

	Query Match	78.6%	Score 2769;	DB 16;	Length 2769;
	Best Local Similarity	100.0%	Pred. No. 0;		
	Matches 2769;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
Qy	338 ATGTCGATAGTGAAAGTTATTATCAAAATTCACACTACTATCAACCTATTCCTAGACT				397
Db	1 ATGTCCGATAGTGAAAGTTATTATCAAAATTCACACTACTATCAACCTATTCCTAGACT				60
Qy	398 GATGAAGTATTGGATGATCATAGAAATCAATCAATCACTAATGATTTGGCAATTAGTAGT				457
Db	61 GATGAAGTATTGGATGATCATAGAAATCAATCAATCACTAATGATTTGGCAATTAGTAGT				120
Qy	458 GAAAGATCATGTTGGATTTAAATCAGATTTGATCAGAAAGTTGTAAAAAGCAAAACAA				517

```
121 GAAGATGAGTTGGAAATTAAATCAGAAATTGAGAAAGTTGTAAGGCAAAAAACA 180
518 CAACAACATCATCAAGAGATTACATCAGTAATGCTAAACCATGACTGTAATCTGCT 577
181 CACACACATCATCAAGAGATTACATCAGTAATGCTAAACCATGACTGTAATCTGCT 240
578 TCTTCAATTAAGAAAAATCTAATCTTACCGATAAAGATGAAATTCACCAACCTATGAGT 637
241 TCTTCAATTAAGAAAAATCTAATCTTACCGATAAAGATGAAATTCACCAACCTATGAGT 300
638 TTATCTGCTGCTGATGATCTAATTAACAGCGCTGCAAAAAATCGTAATTTATTAACCTGAGT 697
301 TTATCTGCTGCTGATGATCTAATTAACAGCGCTGCAAAAAATCGTAATTTATTAACCTGAGT 360
698 TCATTACGTAAGATTTTATTTAATAAGATATATACGACGACAAATTTCTACTAATATCAT 757
361 TCATTACGTAAGATTTTATTTAATAAGATATATACGACGACAAATTTCTACTAATATCAT 420
758 ACTCATCTTGCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 817
421 ACTCATCTTGCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 480
818 AATCAAGAGAAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 877
481 AATCAAGAGAAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 540
878 GGTGCTAATTAATCTAATAATTTTGAATAATTTAGTATGCTGATGCAAAATGAAA 937
541 GGTGCTAATTAATCTAATAATTTTGAATAATTTAGTATGCTGATGCAAAATGAAA 600
938 ACTAATGATAGTAAGATATTAATAATCTAGACCACTGCTAATCATATGAACTTGGT 997
601 ACTAATGATAGTAAGATATTAATAATCTAGACCACTGCTAATCATATGAACTTGGT 660
998 ATTGCTGCTACCAACCTTGTGTGTGGAATGCTGATCTACCGCACTGCTGCTGCTGCT 1057
661 ATTGCTGCTACCAACCTTGTGTGTGGAATGCTGATCTACCGCACTGCTGCTGCTGCT 720
1058 GCTGCTGCTGAGAAAGCAATCTGCTCATCTATTTGATGAGAGCTGATCTGATCATCA 1117
721 GCTGCTGCTGAGAAAGCAATCTGCTCATCTATTTGATGAGAGCTGATCTGATCATCA 780
1118 AGATCATCTCAAGAACTGAGAAAGATGTTTGTCTATGCTGATGCTGATCATCA 1177
781 AGATCATCTCAAGAACTGAGAAAGATGTTTGTCTATGCTGATGCTGATCATCA 840
1178 GTTAATGCAATTTGATTTGATGAAATTTGATGAAATTTGATGAAAGAGAGAGCT 1237
841 GTTAATGCAATTTGATTTGATGAAATTTGATGAAATTTGATGAAAGAGAGAGCT 900
1238 TATTTACAAAAACAATGATGCTAATAATTTTGCCTGATGATGAAATTTCAAAATCTT 1297
901 TATTTACAAAAACAATGATGCTAATAATTTTGCCTGATGATGAAATTTCAAAATCTT 960
1298 TCACAAAAATATATCTAATGCTGATCTGCTCATCTCATCATCATCATCATCATCATCAT 1357
961 TCACAAAAATATATCTAATGCTGATCTGCTCATCTCATCATCATCATCATCATCATCAT 1020
1358 AATTAATAAATAATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1417
1021 AATTAATAAATAATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
1418 CCAAAAAATATTTTAAAGAAAAATTTATGATGATTTTACTCATGATGATTTCTTCA 1477
1081 CCAAAAAATATTTTAAAGAAAAATTTATGATGATTTTACTCATGATGATTTCTTCA 1140
1478 TCTTGAAGAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1537
1141 TCTTGAAGAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
1538 CAATTATCATTAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1597
```

---

```
1201 CAATTATCATTAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
1598 TTGCTGAGCAAGAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1657
1261 TTGCTGAGCAAGAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
1658 CTTTTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1717
1321 CTTTTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
1718 CCAGGTCATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1777
1381 CCAGGTCATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
1778 ACTGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1837
1441 ACTGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
1838 ACTGCTGAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1897
1501 ACTGCTGAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
1898 TATTTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1957
1561 TATTTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
1958 AATAATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2017
1621 AATAATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
2018 TCTCATCGCAAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2077
1681 TCTCATCGCAAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
2078 GCTGATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2137
1741 GCTGATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
2138 CATGGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2197
1801 CATGGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
2198 GATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2257
1861 GATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
2258 AGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2317
1921 AGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
2318 GCTAATTTCTTCTGCTGATTTATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2377
1981 GCTAATTTCTTCTGCTGATTTATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
2378 GCCCACAACAACAACCTATCTATTTTACTTCAACAATTTAATTTAATTTAATTTAATTTAATTT 2437
2041 GCCCACAACAACAACCTATCTATTTTACTTCAACAATTTAATTTAATTTAATTTAATTTAATTT 2100
2438 AATGTTTAAAGAACTTCACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2497
2101 AATGTTTAAAGAACTTCACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
2498 AATCCAACTGAAATATATCTAATATCTAATATCTAATATCTAATATCTAATATCTAATATCTAATATCT 2557
2161 AATCCAACTGAAATATATCTAATATCTAATATCTAATATCTAATATCTAATATCTAATATCTAATATCT 2220
2558 CAACAACAACAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2617
2221 CAACAACAACAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
2618 AGAGCTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2677
2281 AGAGCTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
```

```

QY 2678 AATTATTAGCCTANGAAAAATTTTCAGTCCGTGCAATTCAAATATTAGCTCAATTA 2737
DB 2341 AATTATTAGCCTANGAAAAATTTTCAGTCCGTGCAATTCAAATATTAGCTCAATTA 2400
QY 2738 CAAGTTGAATCAATTCATTCATTAATAAATACCGAAATGTTTCTTAAATTAATTG 2797
DB 2401 CAAGTTGAATCAATTCATTCATTAATAAATACCGAAATGTTTCTTAAATTAATTG 2460
QY 2798 AATGGACAATGTTGTTCCATTAATAATTAGTCAGGGGCTTTTGSTAGTAATGAAGA 2857
DB 2461 AATGGACAATGTTGTTCCATTAATAATTAGTCAGGGGCTTTTGSTAGTAATGAAGA 2520
QY 2858 GTCCCTGGTGAAGTGTACCAATTTAGTGTGTTTTCGAAATGTTGGAATTAATA 2917
DB 2521 GTCCCTGGTGAAGTGTACCAATTTAGTGTGTTTTCGAAATGTTGGAATTAATA 2580
QY 2918 TTTATTAATTATGATTCATTTAATTTGCTCAATGCTGTAATAAATAATGAATAATTC 2977
DB 2581 TTTATTAATTATGATTCATTTAATTTGCTCAATGCTGTAATAAATAATGAATAATTC 2640
QY 2978 AATGAAGACAAAATATGTAATGCAATTTTAAATTCATTCATCAGAAATCAAT 3037
DB 2641 AATGAAGACAAAATATGTAATGCAATTTTAAATTCATTCATCAGAAATCAAT 2700
QY 3038 AGAAGTTAGTTTAAAAAATCATGCTGTAATTAATCAATTAATAGTTTCCCAATAA 3097
DB 2701 AGAAGTTAGTTTAAAAAATCATGCTGTAATTAATCAATTAATAGTTTCCCAATAA 2760
QY 3098 TATGAATAA 3106
DB 2761 TATGAATAA 2769

```

## RESULT 2

```

US-10-128-714-2563
; Sequence 2563, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 2563
; LENGTH: 1992
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-2563

```

```

Query Match 5.1%; Score 179.6; DB 15; Length 1992;
Beet Local Similarity 56.6%; Pred. No. 3e-20;
Matches 376; Conservative 0; Mismatches 279; Indels 9; Gaps 2;
QY 1659 TTTTCATTCGATCAGAAAGAACTATTCATGCCCGCATATTCATCATATGATATCAC 1718

```

```

DB 959 TTTTCATTCGATCAGAAAGAACTATTCATGCCCGCATATTCATCATATGATATCAC 1018
QY 1719 CAGGTCAATCTGTGAGATTAATTAGAAATG-----TGAAGAACTGTGGTGTAG 1772
DB 1019 CTGGGATATCTTCCGGGATCTCTTTCATCTGGCCCGAGGAGGTGTGTGTGTGG 1078
QY 1773 AATGACTATGCTCTACTGATTCGAAATGAATAATGTGCGCAAGATTTGGTATTCATC 1832
DB 1079 AGTACTCAATTCACACAGAGAGAGAAATGTCGACTCTCGGGGATTTTCATCATC 1138
QY 1833 CTTTAATCTGCAAGATTAATTCGAATGCAAGAACTCTGTAAATAAGTTGAATTTAA 1892
DB 1139 CACTACTACGGAAGATATCTTGACTCAAGAAAGCCGTGAAGAGTGTGAGCTTTCAAC 1198
QY 1893 GTTATTAATTTGTTTGTTCATCTT---TGAAGCTATAAGAAATCGAAGATTAAT 1949
DB 1199 AATATTAATCTGTGCTGCTCCGACATTCATCAGCTGACAGACAGACGCTTCA 1258
QY 1950 TAGAACCGATAAATGTTATATGTTGTTTTCATGATGTAATTAACGTTCAATTTT 2009
DB 1259 TGGAGCCCGTCACTTACATGTTGTTTCCGAGACGCTGCTCATTTCTCATTTCA 1318
QY 2010 CACCAATTTCTATCAGCAATGTTAGAAAGAGTTCGTAATTTAGAGATTAATGTCG 2069
DB 1319 CCGAATATCCACACCGCAATGTCAGAGAGAAATGGAAGCTTCGTAATTAATGTC 1378
QY 2070 AATGTAATGCTGATGTTATGTTATGCTTATGATGAATAATTAACGATGTTTGGCC 2129
DB 1379 CCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1438
QY 2130 CCGTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2189
DB 1439 CTGTATCCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1498
QY 2190 GAGATACGATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2249
DB 1499 GCGTCGACGATTTTAAATCTTCTTACCGCAATGCGTGTGTAAGAGTCAATGA 1558
QY 2250 CTTTATGATGATTTATTCAGTAAAGTGAATGATGATGATGATGATGATGATG 2309
DB 1559 GCTTATGATGATTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1618
QY 2310 AAGA 2313
DB 1619 ACGA 1622

```

## RESULT 3

```

US-10-128-714-7563
; Sequence 7563, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362

```

; PRIOR FILING DATE: 2001-08-31  
 ; NUMBER OF SEQ ID NOS: 8603  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 7563  
 ; LENGTH: 1992  
 ; TYPE: DNA  
 ; ORGANISM: Aspergillus fumigatus  
 US-10-128-714-7563

Query Match 5.1%; Score 179.6; DB 15; Length 1992;  
 Best Local Similarity 56.6%; Pred. No. 38-20;  
 Matches 376; Conservative 0; Mismatches 279; Indels 9; Gaps 2;

```

QY 1659 TTTTCATTCTGGAATCAGAGAAATATTCATGCCCCCGATATTCATCATTTAGTATCAG 1718
DB 959 TTTTCTCATCCGAGTCCAGAGACATGTCATGCGCGGAACTGGGGGAACTTGTCCTC 1018
QY 1719 CAGGTCAATCTGTGGAATTTATTTAGAAATG-----TGAAGAAACTGTGTGTTAG 1772
DB 1019 CTGGGATACCTTTCGGGATCTCTTTCACCTTGCCCGAGAGAGGTGTGTGTGTGG 1078
QY 1773 ATTGACTTGTCTCTAGTTCGGAATGAAATGTTGGCCAAAGCAATTTGGTATTCATC 1832
DB 1079 ACGTACTCAATCCAAAGAGAGAGAGTGTGCTGCACTCTCGCGGCAATTTTCATTCATC 1138
QY 1833 CTTTAACTGCTGGAAGATATTCGAATGCAAGAACTGTGAAAGAGTTGAATTTTAA 1892
DB 1139 CACTGACTACGGAAGATATCTGACTCAAGAGCCCGTGAAGAGTGTGAGCTTTTCAAGC 1198
QY 1893 GTTATTTATTTGTGTGTTTCCATCTT---TGAAGCTGATTAAGATCTGAAGTATT 1949
DB 1199 AATATTACTTGTGTGCTTCCGACATCTATCACTCGACAAAGACAGACGACCTTCA 1258
QY 1950 TAGAACCGATTAATGTTATTTGTTTTCATGATGATATTTAAGTTCCATTTT 2009
DB 1259 TGGAGCCCGTCAACTTTCATGATGTTGTTTTCGCGAGCGGTCTCTCATTTCTCATTC 1318
QY 2010 CACCAATTTCTCATCCAGCAAAATGTTAGAGAGAGTTCGTAATTTGAGAGTTATGTCG 2069
DB 1319 CCGAAGATCCACACCGCGCAATGTCAGAGAGAGATTTGGAGAGCTTGTGATTTATG 1378
QY 2070 ATGTTAGTCTGATGTTATTTATTTGCTTAAATGAGAAATTTACCATGTTTTCGCC 2129
DB 1379 CCTGAGTAGTACGTGATCTGTATGTCATATTTGACACATTTGTATTTATTTTGGCC 1438
QY 2130 CCGTATTCATGGAATTTGAATTTGAAGCTGATGCCATTTGAAGATCCGTTTCACTGCTA 2189
DB 1439 CTGTATCCGAGAGATTTGATGATGATCTGAGGCAATTTGAGACCTCGTTTTCATCGCC 1498
QY 2190 GAGATTAAGTATTTAGTATGTTTACAAAGATTTGTTGAATCAAGAGAGAAAGTCAATGA 2249
DB 1499 GCGTGCAGCATTTTGAATCTTCTTACCGCAATCGGTGGCTTGTGTAAGAGTCAATGA 1558
QY 2250 CTTTAAATGAGTTATTTATTCAGTAAAGCTGATGCTATTTAAATGTTTGTATAAAGTGC 2309
DB 1559 GCTTATATGCGCTTCTAGGTGCAAGAGATGTCATTCGCGTTCCTCAAGCGTTTGA 1618
QY 2310 AAGA 2313
DB 1619 ACGA 1622
  
```

## RESULT 4

US-09-864-408A-2265  
 ; Sequence 2265; Application US/09864408A  
 ; Publication No. US20040009474A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Leach, Martin D.  
 ; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Enc  
 ; FILE REFERENCE: 21402-012  
 ; CURRENT APPLICATION NUMBER: US/09/864,408A  
 ; CURRENT FILING DATE: 2001-05-24

; PRIOR APPLICATION NUMBER: 60/206,690  
 ; PRIOR FILING DATE: 2000-05-24  
 ; NUMBER OF SEQ ID NOS: 9068  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2265  
 ; LENGTH: 462  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-864-408A-2265

Query Match 3.3%; Score 118; DB 11; Length 462;  
 Best Local Similarity 67.5%; Pred. No. 3.7e-10;  
 Matches 166; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```

QY 1625 ATTATGAGGTTCAATTAATCAATGATATTTTCACTTTCAATTCATTCGAATCAAGAACT 1684
DB 217 AATACAGAAATTCAAATCTTATGATATTTTGTCTTCCGCTCGAATCTGATGAACT 276
QY 1685 ATTATGCCCCGATATTTCAATCATTTAGTATTCACCAAGTCAATCTGTTGAGATTATT 1744
DB 277 GTGATGCTAGTATATCTCTTGTATATCCGAGGACAGACCTTTATGATGATTATTC 336
QY 1745 AGAATGTTGAAGAACTTGTGTTAGATTGTAATCTGCTACTGATTCGGAATGA 1804
DB 337 AAGGAGCGATCCACATGTTGTTGATTTGCACTTGTCCACTGATGATGAATGCGT 396
QY 1805 ATGTTGCCAAAGCATTTGATTCATCTTTAACTGCTGAAGATTTGAATGCAAGAA 1864
DB 397 TGCAATGCAAAACATTTGAAATCATCTTTTGAACCGGGAAGATATCAGATGCAAGAA 456
QY 1865 ACTGCT 1870
DB 457 ACGCGT 462
  
```

## RESULT 5

US-10-473-126-386  
 ; Sequence 386; Application US/10473126  
 ; Publication No. US20040234973A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Epigenome AG  
 ; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell  
 ; FILE REFERENCE: Proliferative disorders  
 ; CURRENT APPLICATION NUMBER: US/10/473,126  
 ; CURRENT FILING DATE: 2003-09-26  
 ; NUMBER OF SEQ ID NOS: 1258  
 ; SEQ ID NO 386  
 ; LENGTH: 8056  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
 US-10-473-126-386

Query Match 3.3%; Score 117.6; DB 18; Length 8056;  
 Best Local Similarity 42.2%; Pred. No. 1.4e-09;  
 Matches 1426; Conservative 0; Mismatches 1911; Indels 43; Gaps 12;

```

QY 169 TTAATTTAAGATNTTCATTTTATTTTATTTTACCAAGCTATGAATATTTTGTGT 228
DB 1404 TTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 1463
QY 229 CTAAACATATATATATATTTTACAGAAATGCTCAATATTAATTAATTAATTAATTA 288
DB 1464 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1522
QY 289 ATATATTAAGATATATATCCCTTTGTTTATTTTATTTTATTTTATTTTATTTTATTT 348
DB 1523 ATGTTAAATAAATTTTAAATTAATTAATTTTATTTTAAATAATTAATTAATTAAT 1582
QY 349 TGAAGTTATTTATCAAAATTCACATCTATCACTATCTCTATGATCTGATGAAGTAT 408
  
```





; PRIOR APPLICATION NUMBER: US 60/295,890  
 ; PRIOR FILING DATE: 2001-06-05  
 ; PRIOR APPLICATION NUMBER: US 60/303,899  
 ; PRIOR FILING DATE: 2001-07-09  
 ; PRIOR APPLICATION NUMBER: US 60/316,362  
 ; PRIOR FILING DATE: 2001-08-31  
 ; NUMBER OF SEQ ID NOS: 8603  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO: 1563  
 ; LENGTH: 2048  
 ; TYPE: DNA  
 ; ORGANISM: Aspergillus fumigatus  
 ; US-10-128-714-1563

Query Match 3.2%; Score 113.6; DB 15; Length 2048;  
 Best Local Similarity 52.2%; Pred. No. 3.7e-09;  
 Matches 376; Conservative 0; Mismatches 279; Indels 65; Gaps 3;

```

QY 1659 TTTTCATTCTGAATCAGAAAGAACTATTCATGCCCCCATATTCATCATTAAGTATCAC 1718
DB 959 TTTTCATCCAGTCCAGAGACGTGCGAGCGGAGACTGGGGACCTTGCTCC 1018
QY 1719 CAGGTCAATCTGTTGAGATTATTTAGAAATG-----TGAAGAACTGGTGTAG 1772
DB 1019 CTGGTGATCTTCCGGGATCTCTTCACTTGCGCGAGGAGGTGTGTGGTGG 1078
QY 1773 ATGTACTGTCTGCTGATTTGGGAAATGAAATGTGGCCAAAGCATTTGGTATTCATC 1832
DB 1079 AGGTACTCAATCAACAGAGAGAAAGTGTGCACTTCGCGGGCATTTCCATCCATC 1138
QY 1833 CTTTAACGCTGAAGATATTCAGATGCAAGAACTGTAAGAAAGTTGAATTTAA 1892
DB 1139 CACTGACTACGGAAGATATCTGACTCAAGAAAGCCGTAAAGATGCGACTTTCAAC 1198
QY 1893 GTTATTTATTTGTTGTTCCACTT---TGAAGCTGATAAAGATCTGAAGATTAT 1949
DB 1199 AATATTACTTGTCTGCTCCGACATCTATCAGCTGACAAAGACAGACGCTTCA 1258
QY 1950 TAGAACCGATAAATGTTATATGTTGTTTCCAGATGATATTAAGTTCCATTTT 2009
DB 1259 TGAAGCCCGTCAACTTCAATGATGTTGTTCCGACAGGTGTCTCTCATTTCTCAT 1318
QY 2010 CACCAATTTCTCATCAGCAAGATGTTAGAAAGAGTTGTCATATTGAGATTATGTCG 2069
DB 1319 CCGAATATCCACACCGCGAATATGTCAGAAAGATTTGGAAAGCTTGATATGTGT 1378
QY 2070 ATGTAGTGTGATGTTATGCTTAATC-----GATGAAATTTACCGATGTTTGGCCCGGT 2104
DB 1379 CCCTCAGTAGTGAATCTGTATGCGATGATGTAAGTCACAAGGCGGTATTTCTAA 1438
QY 2105 -----GATGAAATTTACCGATGTTTGGCCCGGT 2133
DB 1439 TGAGACCGTCTAGTCTTAACGTTTGCAACAGTACGATGTTGTAAGTTTGGCGCTGT 1498
QY 2134 GATTCATGGAATTTGAATATGAAGTGAATGCAATGCAATGCGTTTCACTGCTAGAGA 2193
DB 1499 CATCCGAGAGATTTGAGATGGAATCTGAGCAATTTGAGACCTGCTTTATGCGCGGT 1558
QY 2194 TACTGATTTTCTAGTATGTTTCAAGAAATGTTGATCAAGAAAGAAAGTCACTT 2253
DB 1559 CGACGATTTTGAATCTTCTTACCGGGAATCGGTGCGCTTGGAAGAAAGTCACTG 1618
QY 2254 AATGAGATTTATATCAGGTAAGCTGATGTCATTAATAATGTTTCTTAAGATGTCAGA 2313
DB 1619 GATGCGCTTCTTAAAGTGGCAAGAGATGTCATGCGGCTTCTCCAAAGCGTTGACAGA 1678
  
```

RESULT 8  
 US-10-128-714-6563  
 ; Sequence 6563; Application US/10128714  
 ; Publication No. US20030119013A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jhang, Bo

; APPLICANT: Hu, Weng  
 ; APPLICANT: Tishkoff, Daniel  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Eroshkin, Alexey M  
 ; APPLICANT: Lemieux, Sebastien M  
 ; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
 ; FILE REFERENCE: 10182-018-999  
 ; CURRENT APPLICATION NUMBER: US/10/128,714  
 ; PRIOR FILING DATE: 2002-04-23  
 ; PRIOR APPLICATION NUMBER: US 60/285,697  
 ; PRIOR FILING DATE: 2001-04-23  
 ; PRIOR APPLICATION NUMBER: US 60/287,066  
 ; PRIOR FILING DATE: 2001-04-27  
 ; PRIOR APPLICATION NUMBER: US 60/295,890  
 ; PRIOR FILING DATE: 2001-06-05  
 ; PRIOR APPLICATION NUMBER: US 60/303,899  
 ; PRIOR FILING DATE: 2001-07-09  
 ; PRIOR APPLICATION NUMBER: US 60/316,362  
 ; PRIOR FILING DATE: 2001-08-31  
 ; NUMBER OF SEQ ID NOS: 8603  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO: 6563  
 ; LENGTH: 2048  
 ; TYPE: DNA  
 ; ORGANISM: Aspergillus fumigatus  
 ; US-10-128-714-6563

Query Match 3.2%; Score 113.6; DB 15; Length 2048;  
 Best Local Similarity 52.2%; Pred. No. 3.7e-09;  
 Matches 376; Conservative 0; Mismatches 279; Indels 65; Gaps 3;

```

QY 1659 TTTTCATTCTGAATCAGAAAGAACTATTCATGCCCCCATATTCATCATTAAGTATCAC 1718
DB 959 TTTTCATCCAGTCCAGAGACGTGCGAGCGGAGACTGGGGACCTTGCTCC 1018
QY 1719 CAGGTCAATCTGTTGAGATTATTTAGAAATG-----TGAAGAACTGGTGTAG 1772
DB 1019 CTGGTGATCTTCCGGGATCTCTTCACTTGCGCGAGGAGGTGTGTGGTGG 1078
QY 1773 ATGTACTGTCTGCTGATTTGGGAAATGAAATGTGGCCAAAGCATTTGGTATTCATC 1832
DB 1079 AGGTACTCAATCAACAGAGAGAAAGTGTGCACTTCGCGGGCATTTCCATCCATC 1138
QY 1833 CTTTAACGCTGAAGATATTCAGATGCAAGAACTGTAAGAAAGTTGAATTTAA 1892
DB 1139 CACTGACTACGGAAGATATCTGACTCAAGAAAGCCGTAAAGATGCGACTTTCAAC 1198
QY 1893 GTTATTTATTTGTTGTTCCACTT---TGAAGCTGATAAAGATCTGAAGATTAT 1949
DB 1199 AATATTACTTGTCTGCTCCGACATCTATCAGCTGACAAAGACAGACGCTTCA 1258
QY 1950 TAGAACCGATAAATGTTATATGTTGTTTCCAGATGATATTAAGTTTCAATTTT 2009
DB 1259 TGAAGCCCGTCAACTTCAATGATGTTGTTCCGACAGGTGTCTCTCATTTCTCAT 1318
QY 2010 CACCAATTTCTCATCAGCAAGATGTTAGAAAGAGTTGTCATATTGAGATTATGTCG 2069
DB 1319 CCGAATATCCACACCGCGAATATGTCAGAAAGATTTGGAAAGCTTGATATGTGT 1378
QY 2070 ATGTAGTGTGATGTTATGCTTAATC-----GATGAAATTTACCGATGTTTGGCCCGGT 2104
DB 1379 CCCTCAGTAGTGAATCTGTATGCGATGATGTAAGTCACAAGGCGGTATTTCTAA 1438
QY 2105 -----GATGAAATTTACCGATGTTTGGCCCGGT 2133
DB 1439 TGAGACCGTCTAGTCTTAACGTTTGCAACAGTACGATGTTGTAAGTTTGGCGCTGT 1498
QY 2134 GATTCATGGAATTTGAATATGAAGTGAATGCAATGCAATGCGTTTCACTGCTAGAGA 2193
DB 1499 CATCCGAGAGATTTGAGATGGAATCTGAGCAATTTGAGACCTGCTTTATGCGCGGT 1558
QY 2194 TACTGATTTTCTAGTATGTTTCAAGAAATGTTGATCAAGAAAGAAAGTCACTT 2253
  
```



```

Db      1559  CGACGATTTGATCTTCTTACCGCGAATCGGTGCTTCTTGAAGAGTCATAGCTT 1618
Qy      2254  AATGAGATTATATACAGGTAAGTGAAGTCAATTAATAATGTTGCTAAAGATGTCAGA 2313
Db      1619  GATGGCCCTTCTTAAAGTGGCAAGAGATGTCTATGCGGCTTCTCCAAAGCTTGCACGA 1678

```

## RESULT 9

```

US-10-128-714-563
; Sequence 563, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroskhin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; PRIOR FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 563
; LENGTH: 4048
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-563

```

```

Query Match      3.2%; Score 113.6; DB 15; Length 4048;
Best Local Similarity 52.2%; Pred. No. 4.9e-09;
Matches 376; Conservative 0; Mismatches 279; Indels 65; Gaps 3;

```

```

Qy      1659  TTTTCATTTCTGAATCAGAGAAGAACTATTCATGCCCCGAGATTCATCATTAAGTATAC 1718
Db      1959  TTTTCTCATCCGAGTCCAGAGCACTGTGCATGCGCGGAAGTGGGGAAGCTTGTCTCC 2018
Qy      1719  CAGGTCAATCTGTTGAGATTATTTAGAAATG-----TGAAGAACTTGTGTGTAG 1772
Db      2019  CTGGTGATACCTTTCGGGATCTTTCAACTTGGCCGAGGAGAGTGTGTGTGTGG 2078
Qy      1773  ATTGTACTTGTCTTACTGATTCGGAATGAAATTTGGCCAAAGCATTGGTATTCATC 1832
Db      2079  ACGTACTCAATCCACAAGAGAGAGAGTGTGCACTCTCGGGCATTTCATCCATC 2138
Qy      1833  CTTTAACTGCTGAAGATATTCGAATGCAAGAACTGTGAAGAAAGTTGAATTAATAA 1892
Db      2139  CACTGACTACGAGAGATATCTTGACTCAAGAGCCCGTGAAGAGTTCAGCTTTCAAGC 2198
Qy      1893  GTTATATATTTGTTGTTTCCATACCTT---TGAAGCTATTAAGATCTGAAGATTAT 1949
Db      2199  AATATTAATTTGTGTCTTCCGGAATCTATCACTGAGCAAGACAGACAGACCTTCA 2258
Qy      1950  TAGAAGCATTAATTTTATATTTGTTGTTTCCATGATGATATTAAGTTCCATTTT 2009
Db      2259  TGGAGCCCGTCAACTTCAATGTGTGTTTCCGAGAGTGTCTCTCATCTTCATTCA 2318
Qy      2010  CACCAATTTTCATCCAGCAATGTTAGAAGAGAGTCTCAATTTAGAGATTAATGTCG 2069
Db      2319  CCGAGATTCACACCGCGCAATGTCAAGAGAGATTTGGAGAGCTTGTATTAATGTGT 2378

```

```

Qy      2070  ATGTTAGTCTGATGTTATGTTATGCTTAATC----- 2104
Db      2379  CCCTCAGTAGAGACGTGATCTGTTATGCAATGATGAATGCAAGAGCGATTTCTTA 2438
Qy      2105  -----GATCAATTAACCATGATGTTTGGCCCGGT 2133
Db      2439  TGAGACCGTCTAGTCTTACAGTTGGCAAGTGAACATGTTGATGTTTGGGCTGT 2498
Qy      2134  GATTCATGAATTTGAATATGAAGCTGATGCCATTTGAAGATCGGTTTCACTGTAGAGA 2193
Db      2499  CATCCGAGAGATTTGATGATGATCTGAGGCAATTGAGACCTCGTTTCAATGCGCGGT 2558
Qy      2194  TACTGATTTTATAGTATGTTTACAAAGATTGTGAATCAAGAGAAAGTTCATGACTTT 2253
Db      2559  CGACGATTTTGAATCTTCTTACCGGAATCGGTGCTTGTGAAGAGTCAATAGCTT 2618
Qy      2254  AATGAGATTATATACAGGTAAGTGAAGTCAATTAATAATGTTGCTAAAGATGTCAGA 2313
Db      2619  GATGGCCCTTCTTAAAGTGGCAAGAGATGTCTATGCGGCTTCTCCAAAGCTTGCACGA 2678

```

## RESULT 10

```

US-10-128-714-5563
; Sequence 5563, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroskhin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; PRIOR FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5563
; LENGTH: 4048
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-5563

```

```

Query Match      3.2%; Score 113.6; DB 15; Length 4048;
Best Local Similarity 52.2%; Pred. No. 4.9e-09;
Matches 376; Conservative 0; Mismatches 279; Indels 65; Gaps 3;

```

```

Qy      1659  TTTTCATTTCTGAATCAGAGAAGAACTATTCATGCCCCGAGATTCATCATTAAGTATAC 1718
Db      1959  TTTTCTCATCCGAGTCCAGAGCACTGTGCATGCGCGGAAGTGGGGAAGCTTGTCTCC 2018
Qy      1719  CAGGTCAATCTGTTGAGATTATTTAGAAATG-----TGAAGAACTTGTGTGTAG 1772
Db      2019  CTGGTGATACCTTTCGGGATCTTTCAACTTGGCCGAGGAGAGTGTGTGTGTGG 2078
Qy      1773  ATTGTACTTGTCTTACTGATTCGGAATGAAATGTTGGCCAAAGCATTGGTATTCATC 1832
Db      2079  ACGTACTCAATCCACAAGAGAGAGATGTGTGCACTCTCGCGGCAATTTTCATCCATC 2138
Qy      1833  CTTTAACTGCTGAAGATATTCGAATGCAAGAACTGTGAAGAAAGTTGAATTAATAA 1892

```



```

Db      2139 CACGACATACCGAAGATATCTTGAACCTCAAGAACCCCGGAAAGGTCCAGCTTTTCAAGC 2198
Qy      1893 GTTATTTTGTGTTGTTCCATCTT---TGAAGCTATAAGAACTGAGATTT 1949
Db      2199 AATATATCTTGTCTGCTTCGAGACATCTATCAGCTGACAGACAGACAGACCTTCA 2258
Qy      1950 TAGAACCGATAAAGTATATATTTGTTTTCATGATGATATATTAAGTTCATTTT 2009
Db      2259 TGGAGCCCGTCAACTTCAACATGTTGTTTCCGAGAGGTGCTCTCATTTCTCATCA 2318
Qy      2010 CACCAATTTCTCATCCAGCAAAATGTTAGAGAGAGTTCATGATGAGATTAATGTCG 2069
Db      2319 CCGAATATCCACACCGCGCAAAATGACAGAAAGAGATTTGGGAAAGCTGATTAATGTCG 2378
Qy      2070 ATGTTAGGCTGATGTTGTTATGTTATGCTTAATC----- 2104
Db      2379 CCTCAGTAGTAGCTGATGATCTGTTATGCCATGATGATTAAGTCAAGAGCGGATTTCTAA 2438
Qy      2105 -----GATGAAATTAACCGATGTTTGGCCCGCT 2133
Db      2439 TGAGACCGCTAGTCTTAACGTGSCACAGTACAGACATTTGTTATGTTTGGGCTGT 2498
Qy      2134 GATTCATGAAATGAAATGAAAGCTGATGCCATTTGAAGATGCCGTTTCTACTGCTAGAGA 2193
Db      2499 CATCCGAGAGATGATGATGATGATGAGGCAATGAGGACCTCGTTTCATCGCGCGGT 2558
Qy      2194 TACTGATTTTGTAGTATGTTTCAAGAAATGTTGTGATCAAGAAAGATGATGATTT 2253
Db      2559 CGAGCATTTTAAATCTTCTTAACCGGAAATCGGTGCGCTTCTGTAAGAGTCTATGAGCTT 2618
Qy      2254 AATGAGATTAATTAATCAGGTAAGCTGATGTCATTAATATGTTGTAAAGATGTCAGAA 2313
Db      2619 GATGCGCTTCTAGGTGCGCAAGACAGATGTCATTCGCGTTTCTCAAGCGTTGCAAGCA 2678

```

## RESULT 11

```

US-10-094-240-10/c
; Sequence 10, Application US/10094240
; Publication No. US20030082637A1
; GENERAL INFORMATION:
; APPLICANT: ZWIBEL, LAURENCE J.
; TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
; FILE REFERENCE: N8289
; CURRENT APPLICATION NUMBER: US/10/094,240
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 10/056,405
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: 60/264,649
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 4985
; TYPE: DNA
; ORGANISM: Anopheles gambiae
US-10-094-240-10

```

```

Query Match      3.0%; Score 106.8; DB 14; Length 4985;
Best Local Similarity 49.5%; Pred. No. 7.4e-08;
Matches 329; Conservative 0; Mismatches 332; Indels 3; Gaps 2;
Qy      345 ATAGGAAAGTTATATCAAAATTCACACTACTCAACCTATTCCTAGATCTGATGAAG 404
Db      3017 ATAAACAACAACAACAACAATAATAGAAATTAATTAATTAACAATGCAATTAATAAGA 2958
Qy      405 TATGATGATCATGAAATCAATCACTAATGATTTGCCATTTAGTGTAGTGAAGATG 464
Db      2957 ATAAATTAACAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATA 2898
Qy      465 AGTTGAAATTAATCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAAT 524
Db      2897 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2838

```

```

Qy      525 ATCATCAAGAGATTACATCAGATTAATGCTAAACCAATTGACTCGTAATCTGGTCTTCAA 584
Db      2837 AATATTAATAA--TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATA 2780
Qy      585 TTAAGAAAAATCTAATCTTACCGATTAAGATTAAGATTAATTAATTAATTAATTAATTAAT 644
Db      2779 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2720
Qy      645 GTGGTG-ATGATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 703
Db      2719 ATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2660
Qy      704 CGTAAGATTTTATTTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 763
Db      2659 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2600
Qy      764 CTGCAATTCATTAATTCATTAATTCATTAATTCATTAATTCATTAATTCATTAATTCATTA 823
Db      2599 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2540
Qy      824 AGAAGAAATCTCAATGGAATTTTACCTCATTAATTAATTAATTAATTAATTAATTAATTAAT 883
Db      2539 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2480
Qy      884 AATATTAATTAATTAATTTGAAATGATTTAGTTAGTCCATGACAAATGAAATGAAATTAAT 943
Db      2479 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2420
Qy      944 GATAGTGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1003
Db      2419 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2360
Qy      1004 GCTA 1007
Db      2359 AATA 2356

```

## RESULT 12

```

US-10-056-405-10/c
; Sequence 10, Application US/10056405
; Publication No. US20030166013A1
; GENERAL INFORMATION:
; APPLICANT: ZWIBEL, LAURENCE J.
; TITLE OF INVENTION: MOSQUITO OLFACTOR GENE, POLYPEPTIDES, AND METHODS OF USE THEREOF
; FILE REFERENCE: N7841
; CURRENT APPLICATION NUMBER: US/10/056,405
; CURRENT FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: 60/264,649
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 4985
; TYPE: DNA
; ORGANISM: Anopheles gambiae
US-10-056-405-10

```

```

Query Match      3.0%; Score 106.8; DB 16; Length 4985;
Best Local Similarity 49.5%; Pred. No. 7.4e-08;
Matches 329; Conservative 0; Mismatches 332; Indels 3; Gaps 2;
Qy      345 ATAGGAAAGTTATATCAAAATTCACACTACTCAACCTATTCCTAGATCTGATGAAG 404
Db      3017 ATAAACAACAACAACAACAATAATAGAAATTAATTAATTAACAATGCAATTAATAAGA 2958
Qy      405 TATGATGATCATGAAATCAATCACTAATGATTTGCCATTTAGTGTAGTGAAGATG 464
Db      2957 ATAAATTAACAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATA 2898
Qy      465 AGTTGAAATTAATCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAAT 524

```





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2005, 14:15:33 ; Search time 10845 Seconds  
(without alignments)  
12372.204 Million cell updates/sec

Title: US-10-018-105-3  
Perfect score: 3525  
Sequence: 1 tataataataataataata.....tcaattatcaatattgc 3525

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1:  
2: gb\_est2:  
3: gb\_hic:  
4: gb\_est3:  
5: gb\_est4:  
6: gb\_est5:  
7: gb\_est6:  
8: gb\_gse1:  
9: gb\_gse2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	359.8	10.2	912	CNS06173	AL399925 T7 end of
2	301.6	8.6	1032	CNS0736W	AL427134 clone BAO
3	211.2	6.0	1089	CNS0603K	AL410742 T3 end of
4	210.8	6.0	798	CNS073G5	AL426171 clone BAO
5	202	5.7	579	BZ295013	BZ295013 CG1236.f1
6	187.2	5.3	911	CNS06H1N	AL399045 T3 end of
7	176	5.0	874	CNS06GM9	AL398239 T3 end of
8	176	5.0	911	CNS06GM7	AL397913 T3 end of
9	172.4	4.9	955	CNS028936	CO028936 EST807320
10	163	4.6	387	CO142543	CO142543 EST837214
11	156.2	4.4	771	CF709049	CF709049 CCAGA79TR
12	146	4.1	973	CF709049	CF709049 T7 end of
13	145.4	4.1	524	BZ297820	BZ297820 CG3653.f1
14	143	4.1	763	AL637812	AL637812 VLS8 CDNA
15	141.2	4.0	614	BM266292	BM266292 VLS8 CDNA
16	141	4.0	712	BQ743184	BQ743184 WHE4101.A
17	133.2	3.8	964	CD034119	CD034119 EST812503
18	128.4	3.6	715	CD488599	CD488599 T11.F12.T
19	122.6	3.5	850	CNS06K53	AL419223 T3 end of
20	121.8	3.5	921	CNS0275P2	AL430413 clone BAO
21	119.2	3.4	907	CO027928	CO027928 EST806312
22	117	3.3	895	CNS06H1J	AL398429 T7 end of
23	114.8	3.3	934	CF823927	CF823927 EST701309
24	110.2	3.1	551	AL637804	AL637804

C 25	109.4	3.1	786	7	CF709038	CF709038 CCAGA79TR
C 26	108.6	3.1	1061	9	CNS07DDE	AL440364 T3 end of
C 27	108.2	3.1	901	7	CO011480	CO011480 EST799815
C 28	108	3.1	965	7	CF820256	CF820256 EST697638
C 29	106.6	3.0	641	8	AQ946120	AQ946120 Sheared D
C 30	106	3.0	575	8	BZ780846	BZ780846 1120H11.9
C 31	102.4	2.9	819	9	CF715298	CF715298 CCAC057F
C 32	102.2	2.9	707	9	CG403581	CG403581 ZMMBB024
C 33	100.6	2.9	815	7	CO025755	CO025755 EST804139
C 34	100.6	2.9	823	7	CO025396	CO025396 EST803780
C 35	100.6	2.9	907	7	CO028288	CO028288 EST806672
C 36	100.6	2.9	932	7	CO025732	CO025732 EST804116
C 37	100.6	2.9	936	7	CO031086	CO031086 EST807319
C 38	100.6	2.9	989	7	CO028935	CO028935 EST807319
C 39	100.4	2.8	413	5	BQ491813	BQ491813 EST00979
C 40	99.8	2.8	587	8	BH873429	BH873429 hp45407.b
C 41	99.8	2.8	781	9	BX145762	BX145762 Danio rer
C 42	97.4	2.8	832	8	BH391984	BH391984 AG-ND-138
C 43	96.2	2.7	772	7	CF816703	CF816703 EST694085
C 44	96	2.7	700	8	AQ940248	AQ940248 Sheared D
C 45	95.4	2.7	670	8	BZ388480	BZ388480 EINDL21TF

## ALIGNMENTS

RESULT 1  
CNS06173/C 912 bp DNA linear GSS 30-NOV-2001  
LOCUS T7 end of clone AS0A018C10 of library AS0A from strain CL18 533  
DEFINITION of Saccharomyces bayanus, genomic survey sequence.  
ACCESSION AL399925  
VERSION AL399925.1 GI:12155087  
KEYWORDS GSS.  
SOURCE Saccharomyces bayanus  
ORGANISM Saccharomyces bayanus  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Saccharomyces.

REFERENCE  
AUTHORS Souci t, J.L., Aigle, M., Artiguenave, F., Blandin, G.,  
Biolotin-Fukuhara, M., Bon, E., Brotier, P., Casaregola, S.,  
de-Montigny, D., Dujon, B., Durrens, P., Lepingle, A., Lorente, B.,  
Malpertuy, A., Neuvigliese, C., Ozier-Kalogeropoulos, O., Potier, S.,  
Saurin, W., Tekala, F., Toffano-Nioche, C., Wesolowski-Louvel, M.,  
Wincker, P. and Weissenbach, J.  
Genomic exploration of the hemiascomycetous yeasts: 1. A set of  
yeast species for molecular evolution studies

TITLE  
JOURNAL FEBS Lett. 487 (1), 3-12 (2000)

PUBMED 1152876  
REFERENCE Bon, E., Neuvigliese, C., Casaregola, S., Artiguenave, F., Wincker, P.,  
Aigle, M. and Durrens, P.  
Genomic exploration of the hemiascomycetous yeasts: 5.  
Saccharomyces bayanus var. uvarum

TITLE  
JOURNAL FEBS Lett. 487 (1), 37-41 (2000)

PUBMED 20584715  
REFERENCE 3 (bases 1 to 912)  
Genoscope.

TITLE  
JOURNAL Direct Submission  
Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,  
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :  
sequencage@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT  
This GSS is part of a random genomic sequencing program of thirteen  
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces  
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,  
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces  
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia  
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,  
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to  
5 kb were prepared and both extremities were sequenced. See  
keywords for description of this sequence and for the sequence of



Oy		1612	AATTCTCATGGAGTATTAATGAGGGTCATTACCTGTATGATTTCACTTTCCATCTCGA	1671
Db		573	AATACCAGAACAAGAACTCCATTGATGACACCATAATGATTTGCCATTTTAGTTCTGA	514
Oy		1672	ATCAGAGAAACTATTTCATGCCCCCAGATATTCATCATTTAGTATGCAGAGTCATCTGT	1731
Db		513	CGTATCTGAACCTTGATGCTCCAGACTCTCGTCACTAGTGAAGCCCTTATCATCTATT	454
Oy		1732	TGCGATTTATTTAGAAAATGCTGAAGAAAATTGGTGGTTAGATTGTACTTGTCTTACTGA	1791
Db		453	TAGACAAATTATTCAATGATGAGTGCTCCATCCGTGGTGGTGAAGCTGCAGATGGCCCTACAGA	394
Oy		1792	TTCCGAATGAAAAATGTTGGGCCAAAGCAATTTGGTATTCATCCTTTAACCTGCGAAGATAT	1851
Db		393	TGAGGAAATGAGGTGCTTGAACCAAACATTTGGAATTCATCCCTTGACATCAGAAAGATAT	334
Oy		1852	TCGATGCAAGAAACTCGTGAAGAAAAGTTGAATTATTATTAAGTTATTATTATTGTTGTTT	1911
Db		333	CAGATGCAAGAAACTCGTGAAGAAAAGTTGAACCTTTTCAAAAACATATTATTCATCTGTTT	274
Oy		1912	CCATACCTTTGAAAGCTGATPAAAAGAAATCTGAAGATTATTAGAACCGATPAATGTTATAT	1971
Db		273	CCATACCTTTGAAAGAAATGATATGAAATCGAGAGACTTTTAGAGCCCAATTAAGTCTATAT	214
Oy		1972	TGTTGTTTTCCATGATAGGTATATTATTAACCTTCATTTTCCACCAATTTCTATCCAGCAA	2031
Db		213	GGTGTTTTTTCGCAACGGTATTCTTAACTTTCCACTTTAACCCAATCTCACATTTGTCCAA	154
Oy		2032	TGTTAGAAGAAGAGTTCGTCAATTGAGAGATTATATGATGTTAGTGTGCTGATGGTTATG	2091
Db		153	TGTCAGAGAGACGTGTGAGACAACTTAGATTTATGTCCGCGTAACTCAGATTTGGATTATG	94
Oy		2092	TTATGCTTAAATCGATGAAATTAACCGATGTTTTTGGCCCCGTGATTCAGAAATGTAATA	2151
Db		93	TTAGCGCTTAGATGATGATATATCACAGATAGTTTTCACCTGTATTATCATCATGATATA	34
Oy		2152	TGAAGCTGATGCCATFGAAGATGCCGTTTCA	2183
Db		33	TGAAGCAGATGCATPAGAAGATTCAGTGTTTA	2
RESULT 3				
CNS06QJK				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
3 (bases 1 to 1089)	Genoscope.	Direct Submission	Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : beqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)	This GSS is part of a random genomic sequencing program of thirteen yeast species: <i>Saccharomyces bayanus</i> var. <i>uvorum</i> , <i>Saccharomyces exiguus</i> , <i>Saccharomyces servazii</i> , <i>Zygosaccharomyces rouxii</i> , <i>Saccharomyces kluyveri</i> , <i>Kluyveromyces thermotolerans</i> , <i>Kluyveromyces lactis</i> var. <i>lactis</i> , <i>Kluyveromyces marxianus</i> var. <i>marxianus</i> , <i>Pichia angusta</i> , <i>Debaryomyces hansenii</i> var. <i>hansenii</i> , <i>Pichia sorbitophila</i> , <i>Candida tropicalis</i> and <i>Yarrowia lipolytica</i> . Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.
FEATURES	source	1. 1089	/organism="Yarrowia lipolytica"	
			/mol_type="genomic DNA"	
			/strain="CLIB 89"	
			/db_xref="taxon:4952"	
			/clone="AM0AA006E08"	
			/clone_1ib="AM0AA"	
			/note="Tend : T3"	
			<3..>833	
			/note="similar to Saccharomyces cerevisiae ORF YFL050c [ ALR2 ; divergent cation transporter ]	
			similar to Saccharomyces cerevisiae ORF YOL130w [ ALR1 ; divergent cation transporter ]"	
			/evidence=not_experimental	
ORIGIN				
Query Match	6.0%; Score 211.2; DB 9; Length 1089;			
Best Local Similarity	61.8%; Pred. No. 3.3e-28;			
Matches 336; Conservative 0; Mismatches 208; Indels 0; Gaps 0;				
QY	1764	GGTGGTTAGATTGTACTTGTCTCTACTGATTCGGAATGAAATGTTGGCCAAACATTTG		18233
Db	1	GGTGGCTGACCTGCTACAAACCCCAACGATGCTGAGTCAAGATCTTTCCAAAGCCTTTG		60
QY	1824	GTATTCATCCTTTAACTGCTGGAAGATTTGGAATGCAAGAACTCGGAAATAATGTAAT		18833
Db	61	GAGTTCACCCCTTGACTTCGAGAGATATCCGAACCCAGAGCTCAGAGAGAGGTGAGA		120
QY	1884	TATTTAAAGTATTTATTTGTTGTTTCCATCTTTTGAAGCTGATAAAGATCTGAG		19433
Db	121	TTTTCAAGTCTACTACTTTGTGTACCTTCAACACTTACGACGAGATTTGGAGTCTGAG		180
QY	1944	ATTATTTAGAACCGATTAATGTTTAAATTTGTTTTCATGATGATATTAACGTTCC		20033
Db	181	ACTACCTTGACCCCTGCTCCCTGTGTACATGCTGTTTTCAAGAGGAGCATCATCTGTC		240
QY	2004	ATTTTTCACCAATTTTCTCATCTCCAGCAATGTTAGAGAGAGTTCGTCATTTGAGAGTT		20633
Db	241	ACTTTTCTCCCTTCCACACCTTCGCTCCCAACGTCACACACAGATCCGACGTGAGAGT		300
QY	2064	ATGTCGATGTTAGTGTGATTTGTTTATGCTTAATGATGAAATTAACGATGGTT		21233
Db	301	ATGTCACGTCACCCCGACCTGATCTGTATGTCATCATTTGATGACATCTGATTCCT		360
QY	2124	TTGCCCCCGGATTCATGAGATTTGAATATGAAAGCTGATGCAATTAAGATGCGTTTCA		21833
Db	361	TCTGCGCCGCTCCCAAGAGATTTGAGTGTGAGCTAGGATTAATGAGAGTCCGTAATTTG		420
QY	2184	CTGCTAGAGATCTGATTTTATAGTAGTATGTTTCAAGAAATTTGGTAATCAAGAGAAAG		22433
Db	421	ATGCTCAGATGACGACTTTTCCCTCATGCTGCGACCAATTTGAAAGCCGATCCAAAG		480
QY	2244	TCATGACTTTAATGAGATTTATCAGGTAAAGCTAGTGCATTAATAATGTTTCTAAA		23033
Db	481	CTGACTTATGATGCTCTTCTTATCCGTTAAAGCTATGATGCTTTCGAAATGTTTCCAAAG		540



**QY** 2304 GATG 2307  
**Db** 541 GATG 544  
**RESULT 4**  
**LOCUS** CNS072G5/c  
**DEFINITION** clone BA0AB011C01 of library BA0AB from strain CLIB 210 of Kluyveromyces lactis, genomic survey sequence.  
**ACCESSION** AL426171.1 GI:12209365  
**VERSION** GSS.  
**KEYWORDS** Kluyveromyces lactis  
**SOURCE** Kluyveromyces lactis  
**ORGANISM** Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetaceae; Saccharomycetaceae; Kluyveromyces.  
**REFERENCE** 1 (bases 1 to 798)  
**AUTHORS** Souciet, J.L., Aigle, M., Artiguenave, F., Blandin, G., Bolotin-Fukuhara, M., Bon, E., Brothier, P., Casaregola, S., de-Montigny, J., Dujon, B., Durrens, P., Leplingle, A., Llorente, B., Malpertuy, A., Neuvéglise, C., Olier-Kalogeropoulos, O., Potier, S., Sautin, M., Tekala, F., Toffano-Nioche, C., Wesolowski-Louvel, M., Winkler, P. and Weissenbach, J.  
**TITLE** Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies  
**JOURNAL** FEMS Lett. 487 (1), 3-12 (2000)  
**MEDLINE** 20584711  
**PUBMED** 11152876  
**REFERENCE** 2 (bases 1 to 798)  
**AUTHORS** Bolotin-Fukuhara, M., Toffano-Nioche, C., Artiguenave, F., Duchateau-Nguyen, G., Lemaire, M., Marnieise, R., Montrocher, R., Robert, C., Termier, M., Winkler, P. and Wesolowski-Louvel, M.  
**TITLE** Genomic exploration of the hemiascomycetous yeasts: 11. Kluyveromyces lactis  
**JOURNAL** FEMS Lett. 487 (1), 66-70 (2000)  
**MEDLINE** 20584721  
**PUBMED** 11152886  
**REFERENCE** 3 (bases 1 to 798)  
**AUTHORS** Genoscope.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (08-SEP-2000) Genoscope - Centre National de Séquençage, 2 rue Gaston Crémieux, CP 5706, 91057 Evry cedex, FRANCE. (E-mail: seqref@genoscope.cns.fr Web: www.genoscope.cns.fr)  
**COMMENT** This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.  
**FEATURES**  
**source** 1..798  
**location/Qualifiers**  
**organism**="Kluyveromyces lactis"  
**mol\_type**="genomic DNA"  
**strain**="CLIB 210"  
**variety**="lactis"  
**db\_xref**="taxon:28985"  
**clone**="BA0AB011C01"  
**clone\_1lb**="BA0AB"  
**complement**("<62..>796")  
**/note**="similar to Saccharomyces cerevisiae ORF YFL050C [ ALR2 ; divalent cation transporter ]  
**/similar** to Saccharomyces cerevisiae ORF YOL130w [ ALR1 ; divalent cation transporter ]"  
**/evidence**="not\_experimental"  
**misc\_feature**  
**ORIGIN**  
**Query Match** 6.0%; Score 210.8; DB 9; Length 798;  
**Best Local Similarity** 72.3%; Pred. No. 3.9e-28;

**Matches** 272; **Conservative** 1; **Mismatches** 103; **Indels** 0; **Gaps** 0;  
**QY** 1948 TTGAAACCGATAATGTTATATGTTGTTTCATGATGATATATTAACGTTCAATT  
**Db** 797 TTGAGCCATTAACGTTATATGTTGTTTGGCCAGCGATTTCACTTCCACTT 738  
**QY** 2008 TTCAACCAATTTCTATCCAGCAATGTTAGAGAGAGTTGTCATTTGAGAGATTATGT  
**Db** 737 TAACCAATCTCAACATTTGTGCAATGTGCAGAGACGTTGAGACAACTTAAGATTATGT 678  
**QY** 2068 CGATGTAGTCGATTTGTTATGTTATGCTTAATGATGATTAACCGATGTTTTC 2127  
**Db** 677 CGCGTAACCTCAATGATGATATATGATGCTTAATGATGATATCAACATATGTTTTC 618  
**QY** 2128 CCCCCTGATTCATGAAATGAAATGAAAGCTGATGCAATGAAAGTCCGTTTCACTGC 2187  
**Db** 617 ACCTGTTATTCATCAATCAATGCAATGATGAGCAATGCAATTAAGATTCAAGTTTATGAC 558  
**QY** 2188 TAGAGATACGATTTTATGATGATGTTTCAAAAGAAATGTCATCAAGAGAAAAGTCAT 2247  
**Db** 557 CAGAGACATGAAATTTGCTGTAATGCTTCAGAGAAATGTTGAGAGAGAGAAAACAAT 498  
**QY** 2248 GACCTTAATGATTAATTAATGAGTAAAGCTGATGCAATTAATGTTGCTAAAGATG 2307  
**Db** 497 GACACTAATGATTAATTAATGAGTAAAGCTGATGCAATTAATGTTGCTAAAGATG 438  
**QY** 2308 TCAAGAGAAAGCTAAT 2323  
**Db** 437 CCAGAGCATATCAAT 422  
**RESULT 5**  
**LOCUS** B2295013/c  
**DEFINITION** CG1236.fl Candida glabrata Random Genomic Library Candida glabrata genomic clone CG1236, genomic survey sequence.  
**ACCESSION** B2295013  
**VERSION** B2295013.1 GI:24436002  
**KEYWORDS** GSS.  
**SOURCE** Candida glabrata  
**ORGANISM** Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetaceae; mitosporic Saccharomycetaceae; Candida.  
**REFERENCE** 1 (bases 1 to 579)  
**AUTHORS** Wong, S., Fares, M.A., Zimmermann, W., Butler, G. and Wolfe, K.H.  
**TITLE** Evidence from comparative genomics for a complete sexual cycle in the 'asexual' pathogenic yeast Candida glabrata  
**JOURNAL** Genome Biol. 4 (2), R10 (2003)  
**MEDLINE** 22508158  
**PUBMED** 12620120  
**COMMENT** Contact: Wong S  
 Department of Genetics, Smurfit Institute  
 Trinity College Dublin  
 Dublin 2, Ireland  
 Tel: 353 1 6082319  
 Fax: 353 1 6798558  
 Email: swong@tcd.ie  
 Class: plasmid ends.  
**FEATURES**  
**source** 1..579  
**location/Qualifiers**  
**organism**="Candida glabrata"  
**mol\_type**="genomic DNA"  
**strain**="CBS 138"  
**db\_xref**="taxon:5478"  
**clone**="CG1236"  
**clone\_1lb**="Candida glabrata Random Genomic Library"  
**ORIGIN**  
**Query Match** 5.7%; Score 202; DB 8; Length 579;  
**Best Local Similarity** 74.9%; Pred. No. 1.6e-26;  
**Matches** 253; **Conservative** 0; **Mismatches** 85; **Indels** 0; **Gaps** 0;

**QY** 1643 CCTGATGATTTTCACCTTTCCATTCGATTCAGAGAAAGCAATTCATGCCCGCATATT 1702

	Dc	344	CCTGACAGATTTTCTTTCTTCATCAGAAATGGAAAGAAGATACATGCACAGATATA	285
OY		1703	CCATCATTTAGTATGACCACGAGTCAATCTGTTCGAGATTATTTAGTAATGTGAAGAAACT	17623
Dd		284	GCAAACCTTAGTGACACCGAGGCAAGTCATCTTTTGACCTTCACAGGGGGGAGAGCCAACG	225
OY		1763	TGTGTGTTAAGATTGTACTCTTGTTCCTACGTATTCGAAATGAAAATGTGGCCAAAGCATTT	18222
Dd		224	TGGTGGCTAGATTGTAGTATCTATCTACATGATGAATGAATGAGATGTCTTGGCAAAGGCTTTTT	165
OY		1823	GGTATTCATCTCTTAACGTGTGAAGATATTTGGAATGCAAGAACTCGTGAAGAAAGTTGAA	18823
Dd		164	GGCATTCATCTTTGACTGTGTGAAGATATTTGAATGCAAGAACTCGTGAAGAAAGTTGAG	105
OY		1883	TTATTTAAAAAGTTATTTATTTTGTGTGTTCCATATCTTTTGAAGCTGATAAAGAACTTGAA	19424
Dd		104	TTATTTAAAGTCTCTACTATTTTGTGTGTTTCCATATCATTTGAAAAATGATTAAGAAATCTGAA	45
OY		1943	GATTATTTTAGAACCGATTAATGTTATATTTGTTGTTT	1980
Dd		44	GATTTCCCTAGAGCCAATCAATGATATATGTGATTTGATT	7
RESULT_6				
CNS06HIN				
LOCUS				
DEFINITION			CNS06HIN	911 bp DNA linear GSS-30-NOV-2001
ACCESSION				T3 end of clone AS04AA012F12 of library AS04A from strain CLIB 533
VERSION				AL399045
KEYWORDS				AL399045.1 GI:12153045
SOURCE				GSS.
ORGANISM				Saccharomyces bayanus
REFERENCE				Saccharomycetes bayanus
AUTHORS				Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
				Saccharomycetales; Saccharomycetaceae; Saccharomyces.
				1 (bases 1 to 911)
				Souciet,U.L., Aigle,M., Artiguenave,F., Blandin,G.,
				Bolotin-Fukuhara,M., Bon,E., Broctier,P., Casaregola,S.,
				de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
				Malpetuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
				Saurin,W., Tekala,F., Toffano-Nicche,C., Wesolowski-Jouvel,M.,
				Wincker,P. and Weissenbach,J.
				Genomic exploration of the hemiascomycetous yeasts : 1. A set of
				yeast species for molecular evolution studies
				FEMS Lett. 487 (1), 3-12 (2000)
				20584711
				2 (bases 1 to 911)
				11552876
JOURNAL MEDLINE				Bon,E., Neuvéglise,C., Casaregola,S., Artiguenave,F., Wincker,P.,
PUBMED				Aigle,M. and Durrens,P.
REFERENCE				Genomic exploration of the hemiascomycetous yeasts : 5.
AUTHORS				Saccharomyces bayanus var. uvarum
TITLE				FEMS Lett. 487 (1), 37-41 (2000)
JOURNAL				20584715
				3 (bases 1 to 911)
				11552880
				Genoscope.
				Direct Submission
				Submitted (07-SEP-2000) Genoscope - Centre National de Séquençage,
				2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
				seque@genoscope.cns.fr Web : www.genoscope.cns.fr)
				This GSS is part of a random genomic sequencing program of thirteenth
				yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
				exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
				Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
				lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
				angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
				Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
				5 kb were prepared and both extremities were sequenced. See
				keywords for description of this sequence and for the sequence of
				the other extremity of this insert.
FEATURES				Location/Qualifiers

```

source
1. .911
/organism="Saccharomyces bayanus"
/mol_type="genomic DNA"
/strain="CLIB 533"
/variety="uvurum"
/db_xref="taxon:4931"
/cclone="AS0AA012P12"
/cclone_lib="AS0AA"
/note="end : T3"
<6. .890
/note="similar to Saccharomyces cerevisiae ORF YOL130w [
ALR1 ; divalent cation transporter ]"
/evidence=not_experimental

ORIGIN
Query Match      5.3%; Score 187.2; DB 9; Length 911;
Best Local Similarity 54.0%; Pred. No. 7,8e-24;
Matches 443; Conservative 1; Mismatches 371; Indels 5; Gaps 3.

QY    1192   TTTTGATGAATTGATGAATTTATTAGAGAAGAAAGAGACGTTATTCAAAAACA 1251
Db     52    TTTCAAAAGTTTTTAAGAGATTTTAAAAAGAAAGAGAAAGTTGGTTTTTAGATGATTT 111

QY    1252   AATGATGTCTAATAATA-TTCTGGTATGTAGTGAATTTCAAATCTTCCAAAATAATA 1310
Db     112   AAAAAAGAAAAAGAGTTAATGCTATAGATTTTGAATGAATGAAGATTAAGTAACT 171

QY    1311   CTACTAGTGGGCATCGCATCATCATATCATCATCAGTAAATATATATAAAAATA 1370
Db     172   TTGGGAGGAGGAAAAAGATTAAGTAATTTGAAAGACATTTAAAGGGGCATGAAAAAT 231

QY    1371   ATGGTGTGATGTGTGTGTTCTAGATATGCCAGATTAAATATATCTCCAAAAATATT 1430
Db     232   TTAGTAAAAAGTTTTAAGATATTGGATTAAAAAGTTTATTTGATTTAAGGTTTTTGG 291

QY    1431   TAAGAATAACATTATCAAGATTGAATTATCTCATGAAAANTCTTCATCTTCAGAGAA 1490
Db     292   AAGCGTTAAGATAAAAAATTAGAGTTTTTGAAGAAVGTGATAGTAAAAATAGAGAA 351

QY    1491   TTTATGAATTGAGACTTAAACAACAACCACCTTCAATATGATGATCAATTATCATPA 1550
Db     352   TAACTGAAAAGAGAGATGAGATGAGATGATGAATTAATTTKGACAAAGTTTGGGGTGT 411

QY    1551   CTTTATCTACATCTTCTACTCTTGAGATCTGGATCTGGGCAAGTGA--TTTGGTGGAG 1607
Db     412   TTTTGTGTAATAAATTAAGTTGAAGGAGAAAGAAAGAAAAATTTAAGATGATGGA 471

QY    1608   CAAGAATTTCTGATGGGATTTAATGAGAGTTCATTACCTGATAGATTTTCACTTTCCATT 1667
Db     472   CATATTTGGTATATGGAATAATATGATTTTTRAGATTTTAAATAGATTTTGAATTTTATGTT 531

QY    1668   CTGAATCAGAGAAGAACATTCATGCCCCCGATATTCATCATCTTATGATCACAGSTCAAT 1727
Db     532   TAGAATATGATGAACCGGTATATGCTATTTGATATTCGATTTNTTGGTATGABAAGTTTATG 591

QY    1728   CTGTTTCAGATTTTATTAGAAATGCTGAAGAACTTGSGTGTGATGTTGATGTTGTCCTA 1787
Db     592   AGTTCTATGAAATGTTTAAAGGGGGAAGAAATTAATATGATGTTGATGAGATTTGCAAG 651

QY    1788   CTGATTCGAAAATGAAAATGTTGGCCAAAGCATTTGGTATTCATCTTTTAACTGCTGAAG 1847
Db     652   CTGATGACGAAATGCCGTGATAGCAAAAGCGTTGCGAAATCCATTCATTCAGACAGAG 711

QY    1848   ATATTCGAATGCAAGAAATCGTGAAGAAAGTGAATTTATTTAAAGTTATATTTTGT 1907
Db     712   ATATTTAGAAATGCAAGAAATCGTGAAGAAAGTGAATTTTCAAGTCAATCAATTTTGT 770

QY    1908   GTTTCCTACTTTTGAAGCTGATTAAGAAATCTGAAGTATTTTGAACCGATTAATGTT 1967
Db     771   GTTTCACACATTCGAAAATGATTTAAGATTCGGAAGATTTCTCGAAGCTATTAACGTT 830

QY    1968   ATATTTGTTTTCATGATGATATTAACGTTCCATTT 2007

```

Db 831 ACATTGTTGTTGATGATCGTGTTTAACTTCATT 870

RESULT 7  
CN506GW 874 bp DNA linear GSS 30-NOV-2001  
LOCUS T3 end of clone AS0AA007H07 of library AS0AA from strain CLIB 533  
DEFINITION of Saccharomyces bayanus, genomic survey sequence.  
ACCESSION AL398239  
VERSION AL398239.1 GI:12151560  
SOURCE GSS  
ORGANISM Saccharomyces bayanus  
Saccharomyces bayanus  
Saccharomyces bayanus  
Bukariyoti; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
REFERENCE Soucier, J. L., Aigle, M., Artiguenave, F., Blandin, G.,  
1 (bases 1 to 874)  
Biolotin-Pukhara, M., Bon, E., Broctier, P., Casaregola, S.,  
de-Montigny, J., Dujon, B., Durrens, P., Lepingle, A., Llorente, B.,  
Malpertuy, A., Neuvéglise, C., Ozier-Kalogeropoulos, O., Potier, S.,  
Saurin, M., Tekala, F., Toffano-Nioche, C., Wesolowski-Louvel, M.,  
Wincker, P. and Weissenbach, J.  
TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of  
yeast species for molecular evolution studies  
JOURNAL FEBS Lett. 487 (1), 3-12 (2000)  
MEDLINE 20584711  
PUBMED 11152876  
REFERENCE 2 (bases 1 to 874)  
Bon, E., Neuvéglise, C., Casaregola, S., Artiguenave, F., Wincker, P.,  
Aigle, M. and Durrens, P.  
TITLE Genomic exploration of the hemiascomycetous yeasts: 5.  
Saccharomyces bayanus var. uvarum  
JOURNAL FEBS Lett. 487 (1), 37-41 (2000)  
MEDLINE 20584715  
PUBMED 11152880  
REFERENCE 3 (bases 1 to 874)  
Genoscope.  
TITLE Direct Submission  
Submitted (07-SEP-2000) Genoscope - Centre National de Séquençage,  
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :  
bege@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
COMMENT This GSS is part of a random genomic sequencing program of thirteen  
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces  
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,  
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces  
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia  
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbophilae,  
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to  
5 kb were prepared and both extremities were sequenced. See  
keywords for description of this sequence and for the sequence of  
the other extremity of this insert.  
FEATURES  
source  
1. 874  
/organism="Saccharomyces bayanus"  
/mol\_type="Genomic DNA"  
/strain="CLIB 533"  
/variety="uvarum"  
/db\_xref="taxon:4931"  
/clone="AS0AA007H07"  
/clone\_11b="AS0AA"  
/note="end : 73"  
misc\_feature  
2. .>715  
/note="similar to Saccharomyces cerevisiae ORF YOL130w [  
ALR1, divalent cation transporter ]"  
/evidence="not\_experimental"

Db 186 CACAACCAAGGGGGAATGATGATTAATCTTAGAGATATCAAGATCATTTGCTGACA 245

RESULT 8  
CN506GW 911 bp DNA linear GSS 30-NOV-2001  
LOCUS T3 end of clone AS0AA005H07 of library AS0AA from strain CLIB 533  
DEFINITION of Saccharomyces bayanus, genomic survey sequence.  
ACCESSION AL397913  
VERSION AL397913.1 GI:12150952  
SOURCE GSS  
ORGANISM Saccharomyces bayanus  
Saccharomyces bayanus  
Saccharomyces bayanus  
Bukariyoti; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
REFERENCE Soucier, J. L., Aigle, M., Artiguenave, F., Blandin, G.,  
1 (bases 1 to 911)  
Biolotin-Pukhara, M., Bon, E., Broctier, P., Casaregola, S.,  
de-Montigny, J., Dujon, B., Durrens, P., Lepingle, A., Llorente, B.,  
Malpertuy, A., Neuvéglise, C., Ozier-Kalogeropoulos, O., Potier, S.,  
Saurin, M., Tekala, F., Toffano-Nioche, C., Wesolowski-Louvel, M.,  
Wincker, P. and Weissenbach, J.  
TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of  
yeast species for molecular evolution studies  
JOURNAL FEBS Lett. 487 (1), 3-12 (2000)  
MEDLINE 20584711  
PUBMED 11152876  
REFERENCE 2 (bases 1 to 911)  
Bon, E., Neuvéglise, C., Casaregola, S., Artiguenave, F., Wincker, P.,  
Aigle, M. and Durrens, P.  
TITLE Genomic exploration of the hemiascomycetous yeasts: 5.  
Saccharomyces bayanus var. uvarum  
JOURNAL FEBS Lett. 487 (1), 37-41 (2000)  
MEDLINE 20584715  
PUBMED 11152880  
REFERENCE 3 (bases 1 to 911)  
Genoscope.  
TITLE Direct Submission  
Submitted (07-SEP-2000) Genoscope - Centre National de Séquençage,  
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :  
bege@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
COMMENT This GSS is part of a random genomic sequencing program of thirteen  
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces  
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,  
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces  
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia  
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbophilae,  
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to  
5 kb were prepared and both extremities were sequenced. See





	RESULT 12
LOCUS	CNS06GSK 973 bp DNA linear GSS 30-NOV-2001
DEFINITION	T7 end of clone ASOAA007A08 of library ASOAA from strain CLIB 533
ACCESSION	ALJ98106
VERSION	ALJ98106.1 GI:12151311
KEYWORDS	GSS.
SOURCE	Saccharomyces bayanus
ORGANISM	Saccharomyces bayanus Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces. 1 (bases 1 to 973)
REFERENCE	Souciat,J.L., Aigle,M., Artiguenave,F., Blandin,G., Boitron-Pichurara,M., Bon,E., Broctier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durieux,P., Lepingle,A., Llorente,B., Maupertuy,A., Neuvéglise,C., Ozier-Kalogiropoulos,O., Potier,S., Saurin,W., Tekala,F., Toffani-Nicchoe,C., Wesolowska-Louvel,M., Wincker,P. and Weissenbach,J. Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies FEBS Lett. 487 (1), 3-12 (2000)
AUTHORS	
TITLE	
JOURNAL	
MEDLINE	20584711
PUBMED	11152876
REFERENCE	2 (bases 1 to 973) Bon,E., Neuvéglise,C., Casaregola,S., Artiguenave,F., Wincker,P., Aigle,M. and Durieux,P. Genomic exploration of the hemiascomycetous yeasts: 5. Saccharomyces bayanus var. uvarum FEBS Lett. 487 (1), 37-41 (2000)
JOURNAL	20584715
MEDLINE	11152880
PUBMED	3 (bases 1 to 973) Genoscope. Direct Submission Submitted (07-SER-2000) Genoscope - Centre National de Séquençage, 2 rue Gaston Crémieu, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : seqref@genoscope.cns.fr Web : www.genoscope.cns.fr)
COMMENT	This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Hansenula lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.
FEATURES	Location/Qualifiers 1..973 /source
ORIGIN	/organism="Saccharomyces bayanus" /mol_type="genomic DNA" /strain="CLIB 533" /variety="uvarum" /db_xref="taxon:4931" /clone="ASOAA007A08" /clone_11d="AS0AA" /note="end : 17" <4..>953 /note="similar to Saccharomyces cerevisiae ORF YOL130w [ Alr1 ; divalent cation transporter ] 1 putative frameshift(s)" /evidence=not_experimental
Query Match	4.1%; Score 146; DB 9; Length 973;
Best Local Similarity	69.8%; Pred.No.2.6e-16;
Matches 208; Conservative	2; Mismatches 87; Indels 1; Gaps 1;
1616 TCTGATGGATTATAGAGGTTCATTACTCGATAGATTTTCACTTTTCATTCCATGCATCA 1675	

[illegible]





```
Qy 1643 CCTGATGATTTTCACCTTTTCATTCGTGAAATCAGAAAGAACTATTCATGCCCCGATATT 1702
Db 199 CCTGACGATTTTGAATTTCTTCATCCGAGGGTGAAGCAAGATCCACGCTCCGAGATT 258
Qy 1703 CCATCATTTAGATCACCAAGTCATCTGTTGAGATTTATTAGAAATGTGAAGAACT 1762
Db 259 GCGGACCTCCATGCGATGATGAGAAATTCACGACCTCTCCGGGGTGGCAAGTGTG 318
Qy 1763 TGGTGGTTAGATGTAATCTTGTCTTCTGATTCGGAATGAAATGTGGCCAAAGCATTT 1822
Db 319 TGGTGGTTGAGATGTTACATCTTACTGAGAAATGGCTATGCTCATGAAAGCTTTT 378
Qy 1823 GGTATTCATCCTTTAACTGCTGAGATATTCGATGCAAGAACTCGTGAAGAAAGTTGAA 1882
Db 379 GCATTCACCCCTTGAACCGCGAGATATTCGGGTTCAAGAGACCCGGAAGAAAGTTGAA 438
Qy 1883 TTATTTAAAGTTATTTATTTTGTGTTTCCATTCCTTTGAAGCTGATPAAAGATCTGAA 1942
Db 439 CTTTCCCGTCACTATTTTGTATGTGCTCCGCTCTTTGTGCAAAATPAGAAATTCGCCA 498
Qy 1943 GATTATTTAGAACCGATAATGTTATATGTTTCCATGATGATATTTAAGTTT 2002
Db 499 GATTTCATGAAACCGTAAATGCTATGTTATGCTTTCCGAGGGGGTGTGAGCTTC 558
Qy 2003 CATTTCACCAATTTCTCATCCAGCAATGTTAGAGAGAGTTCGTCA 2052
Db 559 CACTTTCACCTTGCCCGCACAGTCCCAAGTTAGAGAGCGGATTAGGCA 608
```

Search completed: April 14, 2005, 22:05:42  
Job time : 10852 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: April 14, 2005, 13:38:30 ; Search time 14937 Seconds  
(without alignments)  
11435.003 Million cell updates/sec

Title: US-10-018-105-3  
Perfect score: 3525  
Sequence: 1 tatataataataataata.....tcaattatcacaattatgc 3525

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pac:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	3523	99.9	3525	6	BD273775 Identific
2	3523	99.9	3525	6	AX046919 Sequence
3	2769	78.6	2769	6	AX489375 Sequence
4	2155.2	61.1	2358	6	AR551197 Sequence
5	581	16.5	110000	8	CR382138_14
6	423.8	12.0	507	6	AR551196 Sequence
7	396	11.2	2580	8	AY692762 Saccharom
8	396	11.2	4202	8	SCYO1130W
9	396	11.2	13421	8	SCU41293
10	385	10.9	110000	8	CR380951_1
11	383	10.7	270148	8	YSCCHRVIN
12	376	10.7	110000	8	CR382125_06
13	373.6	10.6	110000	8	AE016817_03
14	263.8	7.5	684	6	CQ446161 Sequence
15	252.4	7.2	35030	8	SPBC27812
16	252.4	7.2	38000	8	AB004539
17	252.4	7.2	38911	8	AB004538
18	215	6.1	110000	8	CR382130_00
19	191.2	5.4	43880	8	SPAC17G6

20	190.8	5.4	110000	8	CR382131_00	CR382131 Yarrowia
21	176.8	5.0	875	11	CNS06F3M	AL395912 T7 end of
22	172.4	4.9	163843	5	BX510939	BX510939 Zebrafish
23	172.4	4.9	211113	2	CR391962	CR391962 Danio rer
24	169.8	4.8	110000	8	CR380959_12	Continuation (13 o
25	168.2	4.8	537	6	A85407	A85407 Sequence 66
26	168.2	4.8	537	6	AR154900	AR154900 Sequence
27	168.2	4.8	537	6	B65425	B65425 Genome DNA
28	150.6	4.3	110000	8	CR382137_08	Continuation (9 of
29	149	4.2	94534	5	AL929250	AL929250 Zebrafish
30	145.4	4.1	173019	5	BX321884	BX321884 Zebrafish
31	145	4.1	238084	2	BX950172	BX950172 Danio rer
32	144	4.1	34347	2	BX957360	BX957360 Danio rer
33	142.8	4.1	155204	2	AC007926	AC007926 Trypanoso
34	142	4.0	177575	5	AL953867	AL953867 Zebrafish
35	141.8	4.0	143342	2	CR388191	CR388191 Danio rer
36	140.2	4.0	145306	2	BX899179	BX899179 Danio rer
37	140.2	4.0	145553	5	AL935272	AL935272 Zebrafish
38	138.2	3.9	154071	3	AC115598	AC115598 Dictyoste
39	130.2	3.7	115990	5	BX890614	BX890614 Zebrafish
40	127.6	3.6	203127	5	BX322566	BX322566 Zebrafish
41	127.4	3.6	3957	8	SCYK1064W	Z28064 S.cerevisia
42	127.4	3.6	28600	8	SCYK1286K	X75781 S.cerevisia
43	126.8	3.6	100893	8	CR380959_13	Continuation (14 o
44	126.2	3.6	146415	5	AL929469	AL929469 Zebrafish
45	126.2	3.6	170800	2	CR352223	CR352223 Danio rer

## ALIGNMENTS

RESULT 1  
LOCUS BD273775 3525 bp DNA linear PAT 17-JUL-2003  
DEFINITION Identification of candida albicans essential fungal specific genes  
and use thereof in antifungal drug discovery.

ACCESSION BD273775.1 GI:33083543  
VERSION JP 2002543799-A/2.  
KEYWORDS  
SOURCE  
ORGANISM

Candida albicans  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; mitosporic Saccharomycetales; Candida.

REFERENCE Roemer T., Bussey H. and Davison J.  
Identification of candida albicans essential fungal specific genes  
and use thereof in antifungal drug discovery  
Patent: JP 2002543799-A 2 24-DEC-2002;  
MCGILL UNIVERSITY

COMMENT OS Candida albicans

PN JP 2002543799-A/2  
PD 24-DEC-2002  
PF 05-MAY-2000 JP 2000616385  
PR 05-MAY-1999 US 60/132878  
PI TERRY ROEMER, HOWARD BUSSEY, JOHN DAVISON  
PC C12N15/09, C07K14/40, C07K16/14, C12Q1/68, G01N33/15, G01N33/50, PC  
G01N33/569,  
PC C12N15/00  
CC Identification of candida albicans essential fungal specific  
genes and use  
CC thereof in antifungal drug discovery  
FH Location/Qualifiers  
FT CDS  
CDS Location/Qualifiers  
1..3525  
/organism="Candida albicans"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:5476"

## FEATURES

source

## ORIGIN

Query Match 99.9%; Score 3523; DB 6; Length 3525;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3525; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible][illegible]

```

Db      2161 TCCGATTGAAGATGCGGTTTTCACCTGCTAGAGATACGATTTAGTAGATGTGTACAAAG 2220
Qy      2221 AATTGGTGAATCAAGAGAGAAAGTCTGCTTTAATAGAGATATATCATGSTRMAAGCTTA 2280
Db      2221 AATTGGTGAATCAAGAGAGAAAGTCTGCTTTAATAGAGATATATCATGSTRMAAGCTTA 2280
Qy      2281 TGTCAATTAATAGTTTGTCTAAAGATGTGAGAGAGAGAGCTAATCTTCTTGTGTTATTA 2340
Db      2281 TGTCAATTAATAGTTTGTCTAAAGATGTGAGAGAGAGCTAATCTTCTTGTGTTATTA 2340
Qy      2341 TCAAGCTCAATATTAATCTTAACAACAACAACAAGGCCCCCAACAACAACAATATCC 2400
Db      2341 TCAAGCTCAATATTAATCTTAACAACAACAACAAGGCCCCCAACAACAACAATATCC 2400
Qy      2401 TATATATCTTCAACCAATTAATTTCAACTTTGATTTAGTTAGAACTTCAACTGG 2460
Db      2401 TATATATCTTCAACCAATTAATTTCAACTTTGATTTAGTTAGAACTTCAACTGG 2460
Qy      2461 TGAAGAGATGAGAGATGAGAGATTAATTTGGTCCCAATCCAACTGAAATTAATACTAA 2520
Db      2461 TGAAGAGATGAGAGATGAGAGATTAATTTGGTCCCAATCCAACTGAAATTAATACTAA 2520
Qy      2521 TACTAATATCTAATACTACTGTTTCACTTCACTCACTCAACAACAACAATGTTAT 2580
Db      2521 TACTAATATCTAATACTACTGTTTCACTTCACTCACTCAACAACAACAATGTTAT 2580
Qy      2581 CACTAACAATCTTTTCCCATCCCGATGACGTCCAAAGCTGATTTGCTATATTT 2640
Db      2581 CACTAACAATCTTTTCCCATCCCGATGACGTCCAAAGCTGATTTGCTATATTT 2640
Qy      2641 AGGTGATTTCAAGATCATATATCAACATGTTTCAAAATTTATTTAGCTTAAGAAAT 2700
Db      2641 AGGTGATTTCAAGATCATATATCAACATGTTTCAAAATTTATTTAGCTTAAGAAAT 2700
Qy      2701 TTTTCACTGCTTCACTCAATTAATTTAGCTCAATTAACAAGTGAATCAATTCACA 2760
Db      2701 TTTTCACTGCTTCACTCAATTAATTTAGCTCAATTAACAAGTGAATCAATTCACA 2760
Qy      2761 TAAATTAATCACCAGAAATGTTTCTAATAATTTACTTTGATTTGGACAATGTTAGTTCATT 2820
Db      2761 TAAATTAATCACCAGAAATGTTTCTAATAATTTACTTTGATTTGGACAATGTTAGTTCATT 2820
Qy      2821 AAATTAATCAGGAGCTTTTGTGTATGATGTAAGTCCCTGAGAGAGTGTACCAA 2880
Db      2821 AAATTAATCAGGAGCTTTTGTGTATGATGTAAGTCCCTGAGAGAGTGTACCAA 2880
Qy      2881 TTTAGGTTGGTTTTCGAAATGTTGAGATTAATTAATTAATTAATTTAGATCAATTTAT 2940
Db      2881 TTTAGGTTGGTTTTCGAAATGTTGAGATTAATTAATTAATTAATTTAGATCAATTTAT 2940
Qy      2941 AATTGCTCAATGTTGTGAAAAAATTTGAATTAATCAATTTGAAGGACAAATTAATGTTAA 3000
Db      2941 AATTGCTCAATGTTGTGAAAAAATTTGAATTAATCAATTTGAAGGACAAATTAATGTTAA 3000
Qy      3001 TGAACCAATTTTAATCAATTCATCAAGAAATCAATTTAGAGGTTTAAAGTTAAAGAA 3060
Db      3001 TGAACCAATTTTAATCAATTCATCAAGAAATCAATTTAGAGGTTTAAAGTTAAAGAA 3060
Qy      3061 TGGTGTATTAATCAATTAATTTAGTTCCCAATTAATTAATTAATTAATTAATTAATTA 3120
Db      3061 TGGTGTATTAATCAATTAATTTAGTTCCCAATTAATTAATTAATTAATTAATTAATTA 3120
Qy      3121 ATGCCACAGAGTTTGAAGTTGTTTTTTTTTTTTTTTTTTTATTTGATGAGAGTTGATAT 3180
Db      3121 ATGCCACAGAGTTTGAAGTTGTTTTTTTTTTTTTTTTTTTATTTGATGAGAGTTGATAT 3180
Qy      3181 AATAATATCTTTTATGAAGTAAACAATAGTAATTAATTAATTAATTAATTAATTAATTA 3240
Db      3181 AATAATATCTTTTATGAAGTAAACAATAGTAATTAATTAATTAATTAATTAATTAATTA 3240
Qy      3241 TTTATATATGATATATATGATGATTAATCTTCTTCTGATTTAGGAGAAAGTTATATTA 3300
Db      3241 TTTATATATGATATATATGATGATTAATCTTCTTCTGATTTAGGAGAAAGTTATATTA 3300

```

```

Db      3241 TTTATATATGATATATATGATGATTAATCTTCTTCTGATTTAGGAGAAAGTTATATTA 3300
Qy      3301 TTTACTATTAACATTAATTTTACAGAGTTGTGTTAAATTTGAGAGTCAATTAATAGAT 3360
Db      3301 TTTACTATTAACATTAATTTTACAGAGTTGTGTTAAATTTGAGAGTCAATTAATAGAT 3360
Qy      3361 GTAAAGAAATTTTAAAGAAATTAAGAAATTAATTAATTAATTAATTAATTAATTAATTA 3420
Db      3361 GTAAAGAAATTTTAAAGAAATTAAGAAATTAATTAATTAATTAATTAATTAATTAATTA 3420
Qy      3421 GGGGGGGGAGAGAAAGGGGATATATATGCGCATTTGTTGTTGTTGTTGTTGTTGTTGTT 3480
Db      3421 GGGGGGGGAGAGAAAGGGGATATATATGCGCATTTGTTGTTGTTGTTGTTGTTGTTGTT 3480
Qy      3481 AAATTAATGTTTATCAATTAATTTATCAATTTATTTATTTATTTATTTATTTATTTG 3525
Db      3481 AAATTAATGTTTATCAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTG 3525

RESULT 2
AX046919
LOCUS      3525 bp      DNA      linear      PAT 15-DEC-2000
DEFINITION      Sequence 3 from Patent WO0068420.
ACCESSION      AX046919
VERSION      AX046919.1 GI:11876349
KEYWORDS
SOURCE
ORGANISM
Candida albicans
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.

REFERENCE
1 Roemer, T., Buessey, H. and Davison, J.
  Identification of candida albicans essential fungal specific genes
  and use thereof in antifungal drug discovery
  Patent: WO 0068420-A 3 16-NOV-2000;
  Mycota Biosciences Inc. (CA)
  Location/Qualifiers
    1..3525
      /organism="Candida albicans"
      /mol_type="unassigned DNA"
      /db_xref="taxon:5476"
      338..33106
        /note="unnamed protein product"
        /codon_start=1
        /transl_table=12
        /protein_id="CAC18924.1"
        /db_xref="GI:11876350"
        /translation="MSDSSEYVQNSTNQPIPRSDVDLDDHRNQITNDCAISDEDEL
        ELKSELSEVVKSEKQOHOHEITSDNAKPLRKSGSSIKKSNLTDKORINPMSEL
        GGDPTINGHKRNRYNMSLRKDFYLKONTDQNSTNNTHLAIPPIPIPTLTAN
        KSRKSCQENLPLIKKTKTIGKNNNNRPENDLVSPMTQKTNDSDDITMTSTANMK
        LGIGATITGEGTATATATTAAGRRRPSRSIDSEADSHARSSEOTBEDYCFPMVG
        DHIRNGIDPDEIDFIREEREAVLQKMLAKNLSRIDEFQNLKNNITSGASHRPY
        HHHSNNKNNNGNDGGSSMALKTPKNIILKTLISREFTHENSSSEIELKTKQ
        OPEYKDOLSLTSSSTSGSGSQVQKGRGISDGINGSLPDRFSLSESESTI
        HAPDIPSLVPGQSVRLDFRNGEETWMDCTCPDSEMKMLAKAGIHLTLEDMO
        ETRKEVELPKSYTFCPTFPEDKSESDLEINVIYVPHGILITFHSRSHRANV
        RRRVRLDLDYDVASDWLCYALIDITQFAVHIGIEADAIEDAVFTARDTDRSS
        MLORIGESRRKVMTLRSGRADYIKMFARKQSEANSSGYORQVNLQOQOQAP
        PPPNPDIITSPINSTLNTSLGTSVGSGVGINFGPPTGNNTNTNTNTGSSBP
        QOQOQHTNKSFPIDARPRADIALYLGIDIDHILITMPONLAYEKLIPRSHSNVLA
        OLQVESFNNSKMTETEMPSKITLIGMLVPLNLYTGLFGMNVVPRGCGTNLMPFGIV
        GVLIPTITIGSPFPAQMWLKKLNNLSIEGQNGNRPIFNHSSRSISLSLGIKGGKNSI
        ISFNPKE"

ORIGIN
Query Match      99.9%; Score 3523; DB 6; Length 3525;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3525; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TATATATATATATATATTAACATTAATTAATTAATTAATTAATTAATTAATTTCACTGTT 60
Db      1 TATATATATATATATATTAACATTAATTAATTAATTAATTAATTAATTAATTTCACTGTT 60

```

61 CAACCCACACCCCTGTGCTAGTCGCAATTTATTAAGTCATATTTGTTGATT 120  
61 CAACCCACACCCCTGTGCTAGTCGCAATTTATTAAGTCATATTTGTTGATT 120  
121 AATTAAACGTCATAGTCCTGAGTCCTGAGTCCTTGTGTTGTTCTATATTAAGA 180  
121 AATTAAACGTCATAGTCCTGAGTCCTGAGTCCTTGTGTTGTTCTATATTAAGA 180  
181 TTTTCATTTTTTTTTTTTACCAAGCTAGGAAATTTATTTTGTGTCTAACACTATA 240  
181 TTTTCATTTTTTTTTTTTACCAAGCTAGGAAATTTATTTTGTGTCTAACACTATA 240  
241 AATAATTTTACGAAATTTGCTACAAATATTAATTAATTAATTAATTAATTAAGA 300  
241 AATAATTTTACGAAATTTGCTACAAATATTAATTAATTAATTAATTAATTAAGA 300  
301 GATATATCCCTTTGTTTTTTTTTTTTCTTCCAGCGATGTCGATAGGAAATTTATA 360  
301 GATATATCCCTTTGTTTTTTTTTTTTCTTCCAGCGATGTCGATAGGAAATTTATA 360  
361 TCAAAATTCAGCTACTAATCAACTATTCCTAGATCTGATGAAGTATGATCATAG 420  
361 TCAAAATTCAGCTACTAATCAACTATTCCTAGATCTGATGAAGTATGATCATAG 420  
421 AATCAATCAGCTAATGATGTTGCTAGTATGATGATGAAGTATGATGATGATGAT 480  
421 AATCAATCAGCTAATGATGTTGCTAGTATGATGATGAAGTATGATGATGATGAT 480  
481 AATAATTAAGTACGAAGTTGTATTAAGGCAAAACAAACAAACAAACAAACAAAC 540  
481 AATAATTAAGTACGAAGTTGTATTAAGGCAAAACAAACAAACAAACAAACAAAC 540  
541 ATCAGATATGCTAAACATGCTGATGATCTGTTCTTCAATTAAGAAAAATCTTA 600  
541 ATCAGATATGCTAAACATGCTGATGATCTGTTCTTCAATTAAGAAAAATCTTA 600  
601 TCTTACCGATTAAGATGATTAACCAACCTTATGATGATGATGATGATGATGAT 660  
601 TCTTACCGATTAAGATGATTAACCAACCTTATGATGATGATGATGATGATGAT 660  
661 TAAACGGGCTCAAAAATCTGATTAATTAACATGATGATGATGATGATGATGAT 720  
661 TAAACGGGCTCAAAAATCTGATTAATTAACATGATGATGATGATGATGATGAT 720  
721 AAAAGTATATGTCAGCAATTTCTATATATCTATCTATCTGCAATTCGATTC 780  
721 AAAAGTATATGTCAGCAATTTCTATATATCTATCTATCTGCAATTCGATTC 780  
781 AATTCGAATTCGAACCCCAATTAATTAATGCTAATTAATTAATTAATTAATTA 840  
781 AATTCGAATTCGAACCCCAATTAATTAATGCTAATTAATTAATTAATTAATTA 840  
841 GGAATTTTACCTCCATTAATTAATTAAGAAACAAATGCTGTAATTAATTAAT 900  
841 GGAATTTTACCTCCATTAATTAATTAAGAAACAAATGCTGTAATTAATTAAT 900  
901 TGAATATGATTTAGTTAGTCCCATGACAAATTAATTAATTAATTAATTAAT 960  
901 TGAATATGATTTAGTTAGTCCCATGACAAATTAATTAATTAATTAATTAAT 960  
961 TAAATACGACCACTGCTAATCATATGAACCTGTGATTTGTTGCTACCAACCT 1020  
961 TAAATACGACCACTGCTAATCATATGAACCTGTGATTTGTTGCTACCAACCT 1020  
1021 TGAAGTGTGATCTACCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
1021 TGAAGTGTGATCTACCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
1081 TGCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
1081 TGCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140

1141 AGATGTTGTTTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
1141 AGATGTTGTTTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
1201 AATTGATGATTTTATTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1260  
1201 AATTGATGATTTTATTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1260  
1261 TAAATATTCGCTATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
1261 TAAATATTCGCTATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
1321 TGCATCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
1321 TGCATCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
1381 TGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
1381 TGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
1441 AATTCAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
1441 AATTCAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
1501 GAAAGCTAAACAAACCACTTACAAATATGATGATGATGATGATGATGATGATGAT 1560  
1501 GAAAGCTAAACAAACCACTTACAAATATGATGATGATGATGATGATGATGATGAT 1560  
1561 ATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620  
1561 ATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620  
1621 TGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
1621 TGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
1681 AACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
1681 AACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
1741 AATTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800  
1741 AATTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800  
1801 GAAATGTTGCGCAAAAGATTTGATGATGATGATGATGATGATGATGATGATGAT 1860  
1801 GAAATGTTGCGCAAAAGATTTGATGATGATGATGATGATGATGATGATGATGAT 1860  
1861 AGAACTGCTGAAAGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
1861 AGAACTGCTGAAAGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
1921 TGAAGCTATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980  
1921 TGAAGCTATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980  
1981 CCATGATGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2040  
1981 CCATGATGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2040  
2041 AAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100  
2041 AAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100  
2101 AATGATGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2160  
2101 AATGATGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2160  
2161 TGCATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220  
2161 TGCATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220  
2221 AATTGATGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2280  
2221 AATTGATGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2280

Db	2221	AAATTGGTAATCCAGAGAAAAGTCAGACTTTAATGAGATTATTAACGTAAGCTGA	2280
Qy	2281	TGTCATTTAAATGTTTGCTAAAGATGCAAGAGAGCTAATTCCTTCTGCTTAATTA	2340
Db	2281	TGTCATTTAAATGTTTGCTAAAGATGCAAGAGAGCTAATTCCTTCTGCTTAATTA	2340
Qy	2341	TCAAGCTCAATATATACCTTACAAACAACAACAAGGCCCCACACACACTAATCC	2400
Db	2341	TCAAGCTCAATATATACCTTACAAACAACAACAAGGCCCCACACACACTAATCC	2400
Qy	2401	TATATATTACTTCACCAATTAATTAACCTTGAATCTTATATAGTTAGAACCTTCAACTG	2460
Db	2401	TATATATTACTTCACCAATTAATTAACCTTGAATCTTATATAGTTAGAACCTTCAACTG	2460
Qy	2461	TGAGAGAGTAGAGAGAGGAATTAATTTTGTCCTCAATCCAACTGAAATTAATCTAA	2520
Db	2461	TGAGAGAGTAGAGAGAGGAATTAATTTTGTCCTCAATCCAACTGAAATTAATCTAA	2520
Qy	2521	TACTATATCTATATCTACTGCTTCACTTCACCACTCAACAACAACAACACTGAT	2580
Db	2521	TACTATATCTATATCTACTGCTTCACTTCACCACTCAACAACAACAACACTGAT	2580
Qy	2581	CACATAACAATCTTTCCCATCCCGAGACAGCTCAAGAGCTATATGCAATATATTT	2640
Db	2581	CACATAACAATCTTTCCCATCCCGAGACAGCTCAAGAGCTATATGCAATATATTT	2640
Qy	2641	AGGTGATATTCAGAGATCAATATATATCAACATGTTTCAAAATTTATTAAGCTATGAAAAAT	2700
Db	2641	AGGTGATATTCAGAGATCAATATATATCAACATGTTTCAAAATTTATTAAGCTATGAAAAAT	2700
Qy	2701	TTTCAGTGTGTCACATTCAAATTAATTTAGTCAATTAACAAGTGAATCAATTCCAA	2760
Db	2701	TTTCAGTGTGTCACATTCAAATTAATTTAGTCAATTAACAAGTGAATCAATTCCAA	2760
Qy	2761	TAAATTAATTCACCGAAATGTTTTCGAAATTTACTTTGAGCAGATGTAATTCATTTCCATT	2820
Db	2761	TAAATTAATTCACCGAAATGTTTTCGAAATTTACTTTGAGCAGATGTAATTCATTTCCATT	2820
Qy	2821	AAATTTAGTCAAGGACCTTTTGGTATGAATGTAAAGAGTCCCTGTGAGAGGTGTACCAA	2880
Db	2821	AAATTTAGTCAAGGACCTTTTGGTATGAATGTAAAGAGTCCCTGTGAGAGGTGTACCAA	2880
Qy	2881	TTTAAAGTTGTTTTGCGAATGTTTGAAGTATTAATATTTAAATTAATTGAGTCAATTTAT	2940
Db	2881	TTTAAAGTTGTTTTGCGAATGTTTGAAGTATTAATATTTAAATTAATTGAGTCAATTTAT	2940
Qy	2941	ATTGCTCAATGAGTGTGTAAGAAAATTTGAATTAATTCATGAGAGCAAAATTAATGTTAA	3000
Db	2941	ATTGCTCAATGAGTGTGTAAGAAAATTTGAATTAATTCATGAGAGCAAAATTAATGTTAA	3000
Qy	3001	TGAGCAAAATTTTATCATTCATCAAGAGATCAATTAGAAGTTTAGGTTTAAAAAAAACA	3060
Db	3001	TGAGCAAAATTTTATCATTCATCAAGAGATCAATTAGAAGTTTAGGTTTAAAAAAAACA	3060
Qy	3061	TGAGTGTATATTAATCAATTAATTAAGTTTCCCAATTAATGAATTAAGATTAATCAAGAA	3120
Db	3061	TGAGTGTATATTAATCAATTAATTAAGTTTCCCAATTAATGAATTAAGATTAATCAAGAA	3120
Qy	3121	ATGCCACAGAGTTGATGCTTTGTTTTTTTTTTTTTTTTTATTTGTCATGATGAGTTGTATAT	3180
Db	3121	ATGCCACAGAGTTGATGCTTTGTTTTTTTTTTTTTTTTTATTTGTCATGATGAGTTGTATAT	3180
Qy	3181	ACATATACTTTTTTATAGAAGTAAACAATGTAAATGAATATAGTGTCTCAATCATCTATA	3240
Db	3181	ACATATACTTTTTTATAGAAGTAAACAATGTAAATGAATATAGTGTCTCAATCATCTATA	3240
Qy	3241	TTTATTAATTTGATATAATCGTATACCTACTCTTGTATTTAGGAGAAAGCTTATATTA	3300
Db	3241	TTTATTAATTTGATATAATCGTATACCTACTCTTGTATTTAGGAGAAAGCTTATATTA	3300
Qy	3301	TTTACTATTAACATTTATTTTTTACAGAGTTGTGTTAAATTTGAGAGTCAATTAATAGGAT	3360
Db	3301	TTTACTATTAACATTTATTTTTTACAGAGTTGTGTTAAATTTGAGAGTCAATTAATAGGAT	3360

Df		3301	TTTACTATTAACCATTTTATTTTTAGGAGTGCTGAATTAAATGGAGAGTCATAAATTAATAGAT	3360
OY		3361	GTAAGAAGAGTTTAAAGAAGAAATTAAGAAATRTATATATAGANGTTCATACGAA	3420
Df		3361	GTAAGAAGAGTTTAAAGAAGAAAGATTAAGAAATATATATATATAGANGTTCATACGAA	3420
OY		3421	GGGGGGGGAAGAGAAAGGAGATATATATCCGCATTTGTGGTACTTTTGTGTTGAATA	3480
Df		3421	GGGGGGGGAAGAGAAAGGAGATATATATCCGCATTTGTGGTACTTTTGTGTTGAATA	3480
OY		3481	AAATATTAAGTTTATCTAAATTTATTAACAATTTATCAATATTTGC	3525
Df		3481	AAATATTAAGTTTATCTAAATTTATTAACAATTTATCAATATTTGC	3525
RESULT 3				
AX489375				
LOCUS		AX489375	2769 bp	DNA linear PAT 16-AUG-2002
DEFINITION			Sequence 6675 from Patent WO02053728.	
ACCESSION		AX489375		
VERSION		AX489375.1	GI:22323387	
KEYWORDS				
SOURCE				
ORGANISM			Candida albicans	
			Candida albicans	
			Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;	
			Saccharomycetales; mitosporic Saccharomycetales; Candida.	
REFERENCE				
AUTHORS		1	Romer,T., Jiang,B., Boone,C., Bussey,H. and Olsen,K.L.	
TITLE			Gene disruption methodologies for drug target discovery	
JOURNAL			Patent: WO 02053728-A 6675 11-JUN-2002;	
			Elittra Pharmaceuticals, Inc. (US)	
FEATURES			Location/Qualifiers	
source		1..2769	/organism="Candida albicans"	
			/mol_type="unassigned DNA"	
			/db_xref="taxon:5476"	
ORIGIN				
Query Match		78.6%; Score 2769; DB 6; Length 2769;		
Best Local Similarity		100.0%; Pwd. No. 0;		
Matches 2769; Conservative		0; Mismatches	0; Indels	0; Gaps
OY		338	ATGTCCGATAGTGAAGTATTTATCAAAATTCACCTACTATCAACCATTCCTAGACT	397
Df		1	ATGTCCGATAGTGAAGTATTTATCAAAATTCACCTACTATCAACCATTCCTAGACT	60
OY		398	GATGAAGATTTGATGATCATAGAAATCCAATCACTATGATGTGGCATTAGTATAGT	457
Df		61	GATGAAGATTTGATGATCATAGAAATCCAATCACTATGATGTGGCATTAGTATAGT	120
OY		458	GAAATGATGTTGAATTTAAATCAGAAATTAGATCAGAAAGTTGTAAAAAGCGAAAACAA	517
Df		121	GAAATGATGTTGAATTTAAATCAGAAATTAGATCAGAAAGTTGTAAAAAGCGAAAACAA	180
OY		518	CAACAACATCATCAAGAGATTACATCAGATATGCTAAACCATGACTCGTAAATCTGGT	577
Df		181	CAACAACATCATCAAGAGATTACATCAGATATGCTAAACCATGACTCGTAAATCTGGT	240
OY		578	TCTTCATTTAAGAAAAATCTAACTTAACCGATTAAGATTAAGATTAACCAACCTATAGT	637
Df		241	TCTTCATTTAAGAAAAATCTAACTTAACCGATTAAGATTAAGATTAACCAACCTATAGT	300
OY		638	TTATCTGGTGTGATGATACATTTAAACAGCGGTCACAAAAATCGTAATTATATACATGAGT	697
Df		301	TTATCTGGTGTGATGATACATTTAAACAGCGGTCACAAAAATCGTAATTATATACATGAGT	360
OY		698	TCATTACGTAAGAATTTTATTTAAAGATATATCTGACGACAAATTCCTACTAATATCAT	757
Df		361	TCATTACGTAAGAATTTTATTTAAAGATATATCTGACGACAAATTCCTACTAATATCAT	420
OY		758	ACTCATCTTGAATTTCAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAAT	817
Df		421	ACTCATCTTGAATTTCAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAAT	480



QY 818 AATCAAGAGAAAAATCTCAATTGAAAAATTTACCTCCATTAATTAAAAAGAAAACAATT 877  
 DB AATCAAGAGAAAAATCTCAATTGAAAAATTTACCTCCATTAATTAAAAAGAAAACAATT 540  
 QY 878 GGTGGTAATTAATTCATTAATTTTGA AAAATGATTTAGTCCCATGCAAAAATGAAA 937  
 DB 541 GGTGGTAATTAATTCATTAATTTTGA AAAATGATTTAGTCCCATGCAAAAATGAAA 600  
 QY 938 ACTAATGATAGTGAAGATATTACTAATAGTACCACTGCTAATCATATGAACTTGGT 997  
 DB 601 ACTAATGATAGTGAAGATATTACTAATAGTACCACTGCTAATCATATGAACTTGGT 660  
 QY 998 ATTGGTCTACAACCCCTTGGTGGTGAAGCTGGTACTACCGCACTGCCACTGCCCTGCT 1057  
 DB 661 ATTGGTCTACAACCCCTTGGTGGTGAAGCTGGTACTACCGCACTGCCACTGCCCTGCT 720  
 QY 1058 GGTGCTGGTGAAGAACATCTGCTCATCTAATTGATAGTGAAGCTGATTCATGATCA 1117  
 DB 721 GGTGCTGGTGAAGAACATCTGCTCATCTAATTGATAGTGAAGCTGATTCATGATCA 780  
 QY 1118 AGATCATCTCAAGAACTGAAGAGATGTTGTTTTCTATGGTGGTGCATCATATTAGA 1177  
 DB 781 AGATCATCTCAAGAACTGAAGAGATGTTGTTTTCTATGGTGGTGCATCATATTAGA 840  
 QY 1178 GTTAAATGAATTTGATTTGATGAAAATTGAATTTTATGAGAAAGAAAGAGAGAGCT 1237  
 DB 841 GTTAAATGAATTTGATTTGATGAAAATTGAATTTTATGAGAAAGAAAGAGAGAGCT 900  
 QY 1238 TATTTTACAAAACAAATGATGCTAATAAATATTCGCGATATGAAATTTGAAAATCTT 1297  
 DB 901 TATTTTACAAAACAAATGATGCTAATAAATATTCGCGATATGAAATTTGAAAATCTT 960  
 QY 1298 TCCAAAAATATATCTACTAGTGGTGCATCTGCTCATCCATCATCATCAAGTAAAT 1357  
 DB 961 TCCAAAAATATATCTACTAGTGGTGCATCTGCTCATCCATCATCATCAAGTAAAT 1020  
 QY 1358 AATTAATAAATAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1417  
 DB 1021 AATTAATAAATAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
 QY 1418 CCAAAAAATATTTTAAAGAAAACATTAATCAAGATTTGAATTTTCAATGAAAATCTTCA 1477  
 DB 1081 CCAAAAAATATTTTAAAGAAAACATTAATCAAGATTTGAATTTTCAATGAAAATCTTCA 1140  
 QY 1478 TCTTGAAGAAATTTATGAAATTTGAAGCTAATAACAACACCTTCAATATGATGAT 1537  
 DB 1141 TCTTGAAGAAATTTATGAAATTTGAAGCTAATAACAACACCTTCAATATGATGAT 1200  
 QY 1538 CAATTATCATTAATCTCATCTAATCTTCTACTCTGATCTGATCTGATCTGATCTGAT 1597  
 DB 1201 CAATTATCATTAATCTCATCTAATCTTCTACTCTGATCTGATCTGATCTGATCTGAT 1260  
 QY 1598 TTTGGTGAAGCAAGAAATTTCTGATGGGATTAATGAGGTTCAATACCTGATGAAATTTTCA 1657  
 DB 1261 TTTGGTGAAGCAAGAAATTTCTGATGGGATTAATGAGGTTCAATACCTGATGAAATTTTCA 1320  
 QY 1658 CTTTTCATCTGAAATCAGAAAGAACTATTCATGCCCCCGATATTCATCATTTAGTATCA 1717  
 DB 1321 CTTTTCATCTGAAATCAGAAAGAACTATTCATGCCCCCGATATTCATCATTTAGTATCA 1380  
 QY 1718 CCAAGTCAATCTGTTGAGATTTATTTAGAAAATGATGAGAAACCTTGGTGGTAAATGTT 1777  
 DB 1381 CCAAGTCAATCTGTTGAGATTTATTTAGAAAATGATGAGAAACCTTGGTGGTAAATGTT 1440  
 QY 1778 ACTGTCCTCACTGATTTGCGAAAATGTAATGTTGGCCAAAGCATTTGGTATTCATCTTTA 1837  
 DB 1441 ACTGTCCTCACTGATTTGCGAAAATGTAATGTTGGCCAAAGCATTTGGTATTCATCTTTA 1500  
 QY 1838 ACTGCTGAAGATATTCGAATGCAAGAACTCGTGAAGAAAGTGAATTTATTTAAAGTTAT 1897  
 DB 1501 ACTGCTGAAGATATTCGAATGCAAGAACTCGTGAAGAAAGTGAATTTATTTAAAGTTAT 1560

QY 1898 TATTTTGTGTTTCCATACCTTTTGAAGCTGATTAAGAAATCTGAAGATTTATTAGAACCG 1957  
 DB 1561 TATTTTGTGTTTCCATACCTTTTGAAGCTGATTAAGAAATCTGAAGATTTATTAGAACCG 1620  
 QY 1958 ATTAATGTTTATATGTTGTTTCCATGATGATGATATTAAGCTTCCATTTTTCACCAAT 2017  
 DB 1621 ATTAATGTTTATATGTTGTTTCCATGATGATGATATTAAGCTTCCATTTTTCACCAAT 1680  
 QY 2018 TCTCATCCAGAAATGTTTGAAGAAAGATTGCTCAATTTGAGAGATTAATGTCATGATAGT 2077  
 DB 1681 TCTCATCCAGAAATGTTTGAAGAAAGATTGCTCAATTTGAGAGATTAATGTCATGATAGT 1740  
 QY 2078 GCTGATTTGTTATGTTATGCTTTAATGATGAATTTACCAGATGTTTGGCCCCGTAT 2137  
 DB 1741 GCTGATTTGTTATGTTATGCTTTAATGATGAATTTACCAGATGTTTGGCCCCGTAT 1800  
 QY 2138 CATGGAATTTGAATPAGAACTGATGCCATGAGAAATGCGTTTCACTGCTAGAGATCT 2197  
 DB 1801 CATGGAATTTGAATPAGAACTGATGCCATGAGAAATGCGTTTCACTGCTAGAGATCT 1860  
 QY 2198 GATTTTATGATAGTTTCAAAAGATTTGGTGAATCAAGAAAGAAAGTCAATGACTTTAATG 2257  
 DB 1861 GATTTTATGATAGTTTCAAAAGATTTGGTGAATCAAGAAAGAAAGTCAATGACTTTAATG 1920  
 QY 2258 AGATTTATTCAGGTAAGCTGATGTCATTAATGTTTCTTAAAGATGTCAGAGAA 2317  
 DB 1921 AGATTTATTCAGGTAAGCTGATGTCATTAATGTTTCTTAAAGATGTCAGAGAA 1980  
 QY 2318 GCTAATCTTCTGCTGTTTATTAATCAAGTCAATTAATCTTAACAACAACAACAACAG 2377  
 DB 1981 GCTAATCTTCTGCTGTTTATTAATCAAGTCAATTAATCTTAACAACAACAACAACAG 2040  
 QY 2378 GCCCCACACCAACCACTAATCTAATTAATTAATTAATTAATTAATTAATTAATTAATCTT 2437  
 DB 2041 GCCCCACACCAACCACTAATCTAATTAATTAATTAATTAATTAATTAATTAATTAATCTT 2100  
 QY 2438 AATAGTTTGAAGATTTCACTGATGAGAGATGAGAGATGAGAGAAATTAATTTGGTCCC 2497  
 DB 2101 AATAGTTTGAAGATTTCACTGATGAGAGATGAGAGATGAGAGAAATTAATTTGGTCCC 2160  
 QY 2498 AATCCAACTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2557  
 DB 2161 AATCCAACTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2220  
 QY 2558 CAACAACAACAACAATGATGATCACTAACAATCTTTCCCATCCCGATGCAAGTCA 2617  
 DB 2221 CAACAACAACAACAATGATGATCACTAACAATCTTTCCCATCCCGATGCAAGTCA 2280  
 QY 2618 AGAGCTGATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2677  
 DB 2281 AGAGCTGATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340  
 QY 2678 AATTTATTAAGCTTAAGAAAATTTTCAAGTGGTGCATTAATTAATTAATTAATTAAT 2737  
 DB 2341 AATTTATTAAGCTTAAGAAAATTTTCAAGTGGTGCATTAATTAATTAATTAATTAAT 2400  
 QY 2738 CAAGTTGAATCAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2797  
 DB 2401 CAAGTTGAATCAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2460  
 QY 2798 ATTGGAGCAATGTTTATGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2857  
 DB 2461 ATTGGAGCAATGTTTATGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2520  
 QY 2858 GTCCCTGGTGAAGGATGAGCAATTTAGGTTGTTTTGGAATTTGTTGAGATTAATA 2917  
 DB 2521 GTCCCTGGTGAAGGATGAGCAATTTAGGTTGTTTTGGAATTTGTTGAGATTAATA 2580  
 QY 2918 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2977  
 DB 2581 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2640  
 QY 2978 ATTGAAGACAAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3037

```

Db      2641 ATTGAAGACAAAATATGATATGACCAATTTTAACTCAATCAAGAGACAAAT
Qy      3038 AGAGTTAGGTTTAAACATGCTGTATTAATCAATATAGTTTCCCAATAA 3097
Db      2701 AGAGTTAGGTTTAAACATGCTGTATTAATCAATATAGTTTCCCAATAA 2760
Qy      3098 TATGATAA 3106
Db      2761 TATGATAA 2769

RESULT 4
ARS51197 2358 bp DNA linear PAT 08-OCT-2004
LOCUS Sequence 6328 from patent US 6747137.
ACCESSION ARS51197
VERSION ARS51197.1 GI:53944372
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 2358)
TITLES Weinstein, K.G. and Bush, D.
Nucleic acid sequences relating to Candida albicans for diagnostics
and therapeutics
JOURNAL Patent: US 6747137-A 6328 08-JUN-2004;
FEATURES Location/Qualifiers
SOURCE 1. 2358
/organism="Unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 61.1%; Score 2155.2; DB 6; Length 2358;
Best Local Similarity 97.2%; Pred. No. 46-292;
Matches 2293; Conservative 0; Mismatches 23; Indels 44; Gaps 8;

Qy      285 ATAAATATTAATTAAGATATATCTCCCTTTGTTTTTTTTTTCTCCAGGCATGTCG 344
Db      1 ATAAATATTAATTAAGATATATCTCCCTTTG--TTTTTTTCTCCAGGCATGTCG 58
Qy      345 ATAGTGAAGTTATATCAAAATCACTACTATCAACTATTCCTAGATCGATGAAG 404
Db      59 ATAGTGAAGTTATATCAAAATCACTACTATCAACTATTCCTAGATCGATGAAG 118
Qy      405 TATTGATGATCATGAATCAATCACTAATGATGTCCTTACTGCTAGTGAAGATG 464
Db      119 TATTGATGATCATGAATCAATCACTAATGATGTCCTTACTGCTAGTGAAGATG 178
Qy      465 AGTTGGAATTAATCAGATTAAGATGAGAGTTGTAAAGGAAA----- 512
Db      179 AGTTGGAATTAATCAGATTAAGATGAGAGTTGTAAAGGAAAAGCAAAACAACAAC 238
Qy      513 AACACAACAACATCATCAAGAGATTACATGATATGCTAAACCTGCTAAAT 572
Db      239 AACACAACAACATCATCAAGAGATTACATGATATGCTAAACCTGCTAAAT 298
Qy      573 CTGCTTTCTTCAATTAAGAAAAATCTAATCTTACCGATTAAGATGAATTACCAACCTA 632
Db      299 CTGCTTTCTTCAATTAAGAAAAATCTAATCTTACCGATTAAGATGAATTACCAACCTA 358
Qy      633 TGAGTTTATTCGCTGTATGATGATTAAGAGGGGTCAAAAATGCTATTTAACA 692
Db      359 TGAGTTTATTCGCTGTATGATGATTAAGAGGGGTCAAAAATGCTATTTAACA 418
Qy      693 TGAGTTCAATTCGTAAGATTTTATTTAAAGATTAATCTAGCACAATTCCTAATA 752
Db      419 TGAGTTCAATTCGTAAGATTTTATTTAAAGATTAATCTAGCACAATTCCTAATA 478
Qy      753 ATCATATCATCTTTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAAT 812
Db      479 ATCATATCATCTTTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAAT 538

```

```

Qy      813 CTAAATAATCAAGAGAAAATCTCAATTTGAAAATTTTACTCTCAATTAATTAAGAAA 872
Db      539 CTAAATAATCAAGAGAAAATCTCAATTTGAAAATTTTACTCTCAATTAATTAAGAAA 598
Qy      873 CAATTGCTGTATTAATTTCTAAATTTTGAATTTGTTAGTATGCTCCCAAAA 932
Db      599 CTATTGGCTGTATTAATTTCTAAATTTTGAATTTGTTAGTATGCTCCCAAAA 658
Qy      933 TGAATACTAATGATGATGATGATTAATA---TACTAGCACCACTCTAATCATATGA 989
Db      659 TGAATACTAATGATGATGATGATTAATAATAATGATTTAGTATGCTCCCAAAA 718
Qy      990 AACTGTATTTGTGTACAACCTTGTGTGAACCTGATCAACCCCACTGCCACTG 1049
Db      729 AACTGTATTTGTGTACAACCTTGTGTGAACCTGATCAACCCCACTGCCACTG 778
Qy      1050 CCACTGCTGCTGC-----TGTGTAAGAACCTCTCTGTTCAATTTGATTAAGAG 1100
Db      779 CCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 838
Qy      1101 CTGATTCATGATCAATGATCATCTCAAGAACTGAAGAGATGTTGTTTCTATG 1160
Db      839 CTGATTCATGATCAATGATCATCTCAAGAACTGAAGAGATGTTGTTTCTATG 898
Qy      1161 TTGCTGATCATATTAAGATTAAGATTTGATGAATTTGATGAATTTATTAAG 1220
Db      899 TTGCTGATCATATTAAGATTAAGATTTGATGAATTTGATGAATTTATTAAG 958
Qy      1221 AAGAAAGAAAGAACTTATTTTCAAAAACAATGATGCTTAAATTTCTGCTATG 1280
Db      959 AAGAAAGAAAGAACTTATTTTCAAAAACAATGATGCTTAAATTTCTGCTATG 1018
Qy      1281 ATGAATTCAAATCTTCCAAAATATTAATCACTAGTGTGATCTGTCATCATATG 1340
Db      1019 ATGAATTCAAATCTTCCAAAATATTAATCACTAGTGTGATCTGTCATCATATG 1078
Qy      1341 ATCATCAAGTATTAATTAATTAATTAATTAATGCTGTATG---GTGCTGTTCTAGTA 1397
Db      1079 ATCATCAAGTATTAATTAATTAATTAATTAATGCTGTATGCTGTATGTTCTAGTA 1138
Qy      1398 TGGCAGCATTAATATATCTCAAAAATATTTTAAAGAAAATTTATCAATTTGAT 1457
Db      1139 TGGCAGCATTAATATATCTCAAAAATATTTTAAAGAAAATTTATCAATTTGAT 1198
Qy      1458 TTAATCAAGAAAATCTTCAATCTGAGAGAAATTTAATGATTAAGAACTAAACAAC 1517
Db      1199 TTAATCAAGAAAATCTTCAATCTGAGAGAAATTTAATGATTAAGAACTAAACAAC 1258
Qy      1518 CACCTTACAATATGATGATCAATTTATCACTTCAATCTTCACT-----T 1571
Db      1259 CACCTTACAATATGATGATCAATTTATCACTTCAATCTTCACTCTTCTGAT 1318
Qy      1572 CTGATCTGATCTGGGCAAGTGAATTTGTGAGCAAGAAATTTCTGATGGATTAATG 1631
Db      1319 CTGATCTGATCTGGGCAAGTGAATTTGTGAGCAAGAAATTTCTGATGGATTAATG 1378
Qy      1632 GAGGTTCAATTAATGATGATTTTCACTTCCATTCGAATCAGAGAACTATTCATG 1691
Db      1379 GAGGTTCAATTAATGATGATTTTCACTTCCATTCGAATCAGAGAACTATTCATG 1438
Qy      1692 CCCCCGATTAATCATTAATGATCAACAGGTCATCTGTTGAGATTTATTAAGATG 1751
Db      1439 CCCCCGATTAATCATTAATGATCAACAGGTCATCTGTTGAGATTTATTAAGATG 1498
Qy      1752 GTGAAGAACTTGTGTGATGATTTCTTGTCTGATGTTGGAATGAAGATTTGG 1811
Db      1499 GTGAAGAACTTGTGTGATGATTTCTTGTCTGATGTTGGAATGAAGATTTGG 1558
Qy      1812 CCAAGCAATTTGATTTGATCTTAACTGCGAAGATTTGATGCAAGCAAGCACTGG 1871
Db      1559 CCAAGCAATTTGATTTGATCTTAACTGCGAAGATTTGATGCAAGCAAGCACTGG 1618
Qy      1872 AAAAGTTGAATTAATTAAGATTAATTTTGTGTTCTCATCTTTGAGAGCTGATA 1931

```

```

Db      1619  AAAAGTGAATATATTAAGCTATATTTGTTGTTCCATCTTTGAGCGTGA 1678
Qy      1932  AAGATCTGAAGATATTAGAACCGATTAATGTTATATGTTGTTCCATGATGTA 1991
Db      1679  AAGATCTGAAGATATTAGAACCGATTAATGTTATATGTTGTTCCATGATGTA 1738
Qy      1992  TATTAAGCTTCAATTTTCCCAATTTCTCATCGAACAAATGTTAGAAAGTTCGTC 2051
Db      1729  TATTAAGCTTCAATTTTCCCAATTTCTCATCGAACAAATGTTAGAAAGTTCGTC 1798
Qy      2052  AATTAGAGATTAATGCTCATGTTAGTGTGATTTGTTATGTTATGCTTAATCGATGAA 2111
Db      1799  AATTAGAGATTAATGCTCATGTTAGTGTGATTTGTTATGTTATGCTTAATCGATGAA 1858
Qy      2112  TTACCGATGTTTTCGCGCGGATTCATGAAATGAATATGAGCTGATGCTGATGAG 2171
Db      1859  TTACCGATGTTTTCGCGCGGATTCATGAAATGAATATGAGCTGATGCTGATGAG 1918
Qy      2172  ATGCCGTTTCACTGCTAGATGATGATTTTATGTTATGTTACAAAGAAATGCTGAT 2231
Db      1919  ATGCCGTTTCACTGCTAGATGATGATTTTATGTTATGTTACAAAGAAATGCTGAT 1978
Qy      2232  CAAGAGAAAGTCTGATGATTTATGATTAATATGAGTAAAGCTGATGCTTAATAA 2291
Db      1979  CAAGAGAAAGTCTGATGATTTATGATTAATATGAGTAAAGCTGATGCTTAATAA 2038
Qy      2292  TGTGTTGTTAAAGATGTCAGAGAAAGCTAATTTCTTCTGTTTATTTATCAAGCTCAT 2351
Db      2039  TGTGTTGTTAAAGATGTCAGAGAAAGCTAATTTCTTCTGTTTATTTATCAAGCTCAT 2098
Qy      2352  ATAACTTCAACAACAACAACA-----GGCCCAACAACAACAACAACAACAACAACA 2405
Db      2099  ATAACTTCAACAACAACAACAACAAGGCCCAACAACAACAACAACAACAACAACA 2158
Qy      2406  TTACTTCAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2465
Db      2159  TTACTTCAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2218
Qy      2466  GAGTAGAGTAGAGAGAAATTAATTTTGGTCCCAATCGAATGGAATATATATATATCA 2525
Db      2219  GAGTAGAGTAGAGAGAAATTAATTTTGGTCCCAATCGAATGGAATATATATATATCA 2278
Qy      2526  ATACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2585
Db      2279  ATACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2335
Qy      2586  ACAAATCTTTCCCAATCCCC 2605
Db      2336  ACAAATCTTTCCCAATCCCC 2355

```

## RESULT 5

CR382138\_14/c

## WPCOMMENT

Sequence split into 24 fragments LOCUS CR382138 Accession CR382138

```

Fragment Name      Begin      End
CR382138_00      1      110000
CR382138_01      100001    210000
CR382138_02      200001    310000
CR382138_03      300001    410000
CR382138_04      400001    510000
CR382138_05      500001    610000
CR382138_06      600001    710000
CR382138_07      700001    810000
CR382138_08      800001    910000
CR382138_09      900001   1010000
CR382138_10     1000001   1110000
CR382138_11     1100001   1210000
CR382138_12     1200001   1310000
CR382138_13     1300001   1410000
CR382138_14     1400001   1510000
CR382138_15     1500001   1610000

```

```

CR382138_16      1600001   1710000
CR382138_17      1700001   1810000
CR382138_18      1800001   1910000
CR382138_19      1900001   2010000
CR382138_20      2000001   2110000
CR382138_21      2100001   2210000
CR382138_22      2200001   2310000
CR382138_23      2300001   2336804
Continuation (15 of 24) of CR382138 from base 1400001 (CR382138 Debaryomyces hansenii chr
Query Match      16.5%; Score 581; DB 8; Length 110000;
Best Local Similarity 58.1%; Pred. No. 1.9e-72;
Matches 1348; Conservative 0; Mismatches 815; Indels 158; Gaps 12;
1067  AAGAGACATCTCGTTCATCTATGATGATGAGCTGATTCATGATCAATCATCT 1126
Db      14194  AAGAGACATCTCGTTCATCTATGATGATGAGCTGATTCATGATCAATCATCTCA 14135
Qy      1127  CAAGAACTGAAGAGATGTTGTTTCTATGTTGTTGATCATATTAAGTTAATGGA 1186
Db      14134  CAGGAGACAGAAAGATGTTGCTTCTATGTTGCTGAAACGTTGTTAAGGCT 14075
Qy      1187  ATGATTTGATGAAATGATGAATTTATAGAG--AGAAAGAGAGAGAGCTTATTAC 1244
Db      14074  ATGATTTGATGAAATGATGAAGTTCAATTAAGATGAGAAAGAGAAATGATCATTA 14015
Qy      1245  AAAACAATGATGCTAATAATATTCGCGTATGATGAATTTCAAAATCTTCCAAA 1304
Db      14014  AAAAGAGAACACAGATGATGCGAAGAAAGAACGCGATGAGGCT-----TGATCTGAC 13960
Qy      1305  ATAACTACTAGTGTGATCTGTCATCATCATATCATCATCATCATATTAATAATA 1364
Db      13959  ATGGAGTGCAGCGATGATGATGTC-CCAGTAAATATGCTTGAATATATCTCCGCAAC 13901
Qy      1365  AAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1424
Db      13900  ATATTAGAAAGGCTTCGTTGGAAGTTGTTGAAAGAAATGATGATGATGATGATGAT 13841
Qy      1425  ATATTTTAAAGAAACATTAATCAAGATTTGAATTAATCAATTAATCAATTAATCA 1484
Db      13840  ATGCTTCAATCTGATTAATTAAT-ATTTGAGAGAGGCAATCAATCAATCTTCAACGTTA 13783
Qy      1485  AAGAAATTAATGAATGGAAGCTAACAACAACAACCTTAACAATATGATGATCAATTA 1544
Db      13782  CAGATGATTTGCTCTGAGATGATCAACAAGAAATTTGGAGACATATATTTGGA 13723
Qy      1545  CATTAATCTTCAATCAATCTTCTAATCTTCTAATCTTCTAATCTTCTAATCTTCTA 1604
Db      13722  CGAAA-----ATGTTAAGTTTGGTG 13702
Qy      1605  GAGCAGAAATTTCTGATGATTAATGAGGTTCAATCAATGATGATGATGATGATGAT 1664
Db      13701  GTACAGAAATCAACA-----TAATGATGAGGATTTGCTGAAAGTTTCAATCTTTT 13648
Qy      1665  ATTCTGAATCAAGAAAGTATTCATGCGCCCGATATTCATCAATTAATCAACAGTTC 1724
Db      13647  GTTCGGAATCAAGAAAGTATTCATGCTCTGATATCTTCTGTTGTTCAAACTCGGCG 13588
Qy      1725  AATCTGTCGATTAATTAATGAATGTTGAAGAACTTGTGTTGATTAATTAATTAAT 1784
Db      13587  AGAGTAGAGGATTTGTTCAAGAAATGTTGAAGGATCAATGATGATGATGATGATGAT 13528
Qy      1785  CTACTGATTCGAAATGAAGATGTTGCGCAAGCATTTGATTAATCAATCTTTAATGCTG 1844
Db      13527  CTACTGATTCGAAATGAAGATGTTGCTTAAGCATTTGATTAATCAATCTTTGACAGCTG 13468
Qy      1845  AAGATATTGAAATGCAAGAACTCGTGAAGAAAGTTGAATTAATTAAGTTAATTTTG 1904
Db      13467  AAGATATTGATGCAAGAAACCGTGAAGAAAGTGAATTAATTAAGAAATTAATTTTG 13408
Qy      1905  TTTGTTTCCATTAATTTGAAGCTGATTAAGAAATGATTAATTAATTAATTAATTA 1964
Db      13407  TTTGTTTCCATTAATTTGAAGGATGATGAGAAATCAAGAGATTAATTAATTAATTA 13348

```

Oy	1965	TTTATATATGTTGTTTCCAGAGATATTAATTAACGTTCCATTTTTCACCAATTTCTCAC	2024
Db	13347	TTTATATGTTGTTTCCAGAGAGGATATTTGCTGTTTCATTTTTCGCTATTTTATATC	132888
Oy	2025	CAGCAATGTTAGAGAAAGTTTCGTCAATTGAGATATATGTCATGTTAGTCTGATT	2084
Db	13287	CAGCCAAATGTCAGAAAGAAAGTCCGTCAATTAAAGATTATCGTTAGTCAAGTCCGATT	132287
Oy	2085	GGTATATGTTATGCTTATATGATGAAATTCGATGTTTTGGCCCCGGATTCATGAA	2144
Db	13227	GGATCTGTTATGCTTATATTTGAATGACATTACATATGATTCGCCCCCTGTTATTAACGTG	131687
Oy	2145	TTGAAATATGAAGCTGATGCAATTGAAAGATCGGTTTTCACTGCTAGATATCATGATTTTA	2204
Db	13167	TCGAATATGAAGCTGATGCAATGCAGATATTCGATTTGCTGTAAACAAGTTGGTTTTA	131087
Oy	2205	GTAATATGTTATCAAGAAATTTGGTGAATCAAGAAAGAAAGTCATGACTTTTATATGATATAT	2264
Db	13107	GTACAAATGATATCAAGAGATTTGGTGAATCAAGAAAGAAAGTATATGACTTTGATGATATAT	130487
Oy	2265	TATCAGGTAAAGCTGATGTCATTAAATTTGTTGCTAAAGATGTCAAGAAAGAC-----	2319
Db	13047	TGTCGGAAAAAGCGAGGTTATCAAAATGTTTGTAAAGAATGTCAAGATGAGGCTGCA	129887
Oy	2320	-----TATTTCTTCTCTGTTTATATCAAGCTCAATATATATCTTACAAACA---	2367
Db	12987	ATTAACAAATTAATAATATATTCGAGGCATATCCAAACATGAAACGAAAGGGTAAATTAATAT	129287
Oy	2368	-ACAAACAAGGCCCAACAACAACCACTATATCTTATATCTTACCAAAATTAATTCAA	2426
Db	12927	CACAAACAAGGGTAAATGCTTCCGAAGCCATACATATGATATATACACATGTT	128687
Oy	2427	CTTTGAATCTTAATATGTTTAGAACTTTCACATGCTGAGAGAGATGAGAGAAATTA	2486
Db	12867	ACATTAACGAAGCCCTTACCAACAAGTATTAACAATATTCAAACCAACGAGGCTATA	128087
Oy	2487	ATTTTGGTCCCATTCCACTGGAATATATCTAATATCTAATATCTAATATCTACTGTTCA	2546
Db	12807	ACAAATATTTCAAAACCAAGGTTATATCAATATATTCAAACCAAGGTTATAGCAATA	127487
Oy	2547	CTTACCAACCTCAACAACAACAACAATGATATCACTAACATCTTTCCCATCCCG	2606
Db	12747	ATCCAAACCAACAAGATTAATATCAAGCGGCGCCAGGTGTAAGTTTGGCCAGGTCAG	126887
Oy	2607	ATG-----	2609
Db	12687	GTATATCTAACACATATGATCCGGGTGATTTTATATGAAGATCAACCCAGATAGAA	126287
Oy	2610	CAGCTCCAGAAGCTGATATTCATTATATTTAGGTGATATTCAGATCATATATATCACA	2669
Db	12627	CTCAACCAAGGGCTGACATAGCGCTATATTTGGGCGATATTCAGATCAAGTGTACTTA	125687
Oy	2670	TGTTTCAAAATTTATAGCTTATGAAAAAAATTTTCAGTGTTCACATTCAAATATTAG	2729
Db	12567	TGTTTCAAAATTTATAGTCTTACGAAAGATTTTATGATGCTCCATTCAAACTATTG	125087
Oy	2730	CTCAATTAACAAGTGAATCATTCATTCGAATATATAAATCAACGAAATGTTTTCTTAAA	2789
Db	12507	CACATTAACAAGTGAATCTTATATTTTAATATATAGTTACTGTCTATGTTATCTTAATG	124487
Oy	2790	TTACTTGAATTTGGACATGTTATGTTCCATTAAATTTAGTCAACGGAATCTTTTGGATTGA	2849
Db	12447	TCAAAATGATATGATAGGTCCTTTTCCATTGAATGATCACTGGGTTATTTTGGATTGA	123887
Oy	2850	ATGTAAGATCTCTGTGGAAGGTGT-----ACCAATTAAGTGTGTTTTCGAATTG	2903
Db	12387	ACGTGCTGTGTTCCAGGTCAACAGGTATATCAACCAAGTCTTGTGCTGTTCTTCGATATG	123287
Oy	2904	TTGAGATATTAATATTTATATATTTATGATCATTTAATTTGCTCAATGCTGTGTATATA	2963
Db	12327	TTGGGATATGTGTGCTAATGTTATATGTGTGTAGTTTTCGACGATATTTGATCTGCA	122687

Oy	AATGATTAATTCATTTGAAGACAAATATGTATGACCAATTTTAAATCATTCAT	3023
Db	12267 GAGTCACGGAAAGAACCTATGATGAGAAAACCAAAAAATTCATTTAAAGTAGG	12208
Oy	3024 CAGGAGATCAATTAGAGTTTAGTTTAAAAAAAATGATGATTAATAATCAATTATTA	3083
Db	12207 TTAATA-----CGAGACAAAAGAGCTGCAAAATGATATATCA	12166
Oy	3084 GTTTCGCCAATPAAATATGATTAAGATTAATCAAGAAATGCCACAGAGTTTGATG	3143
Db	12168 GCCTCCCAATTAAGTAGCATTA--TATAAGATGCTAATTACAAATTAATTAATCTTTTG	12110
Oy	3144 TTTTTTTTTTTTTATGTGATGATGAGAGTGTATATACATATCTTTTATGAGTAA	3203
Db	12109 CATATGATTCATTAATTTTACAGAGATTAGATTAATTAATTAATATAGTCTAATGCA	12056
Oy	3204 CAATAGTAATGATPAAATAGATGCATCATCATCATTTATTAATGTATATAATGCTAT	3263
Db	12049 TCAACATTTTGGATTTGTATCGCCCCGATCTGAATGAGATTTTAAAGCTGTGGC	11996
Oy	3264 ACTAATCTTTCTTGATTTTGGGAAAGAGTATATATTATTA 3304	
Db	11989 TTTTAGTTGCATTTTCTTATTAATTAATTAAGTATATCTCTTATA 11949	

RESULT 6	ARS51196	507 bp	DNA	Linear	PAT 08-OCT-2004
LOCUS	ARS51196				
DEFINITION	Sequence 6327 from patent US 6747137.				
ACCESSION	ARS51196				
VERSION	ARS51196.1	GI:53944371			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 507)				
TITLE	Weinstock,K.G. and Bush,D.				
	Nucleic acid sequences relating to Candida albicans for diagnostics and therapeutics				
	Patent: US 6747137-A 6327 08-JUN-2004;				
JOURNAL	Location/Qualifiers				
FEATURES	1..507				
Source	/organism="unknown"				
	/mol_type="genomic DNA"				
ORIGIN					
	Query Match	12.0%;	Score 423.8;	DB 6;	Length 507;
	Best Local Similarity	98.0%;	Pred. No. 8.3e-50;		
	Matches 440;	Conservative 0;	Mismatches 7;	Indels 2;	Gaps 1;
QY	2660 ATATCCACCAATGTTCCAAATTTATTTG--CCATATGAAAAAATTTTCAGTGCACATT	2717			
DB	59 AGACTCACCAATGTTCCAAATTTATTTAGGCTTAATGAAAAAATTTTCAGTGCACATT	118			
QY	2718 CAAATTTATTTAGTCAATTAACAAGTGAATCATTCATTCCAATATATTAATACACCGAA	2777			
DB	119 CAAATTTATTTAGTCAATTAACAAGTGAATCATTCATTCCAATATATTAATACACCGAA	178			
QY	2778 TGTTTTCTAAAAATTTACTGTTGATTTGGGCAATGTTAGTTCATTAAATTTAGTCACGGAC	2837			
DB	179 TGTTTTCTAAAAATTTACTGTTGATTTGGGCAATGTTAGTTCATTAAATTTAGTCACGGAC	238			
QY	2838 TTTTGTGATGCAATGTAAAGAGTCCCTGTGTAAGTGTGATCCAAATTTAGTGTGTTTTCG	2897			
DB	239 TTTTGTGATGCAATGTAAAGAGTCCCTGTGTAAGTGTGAAAAATTTAGTGTGTTTTCG	298			
QY	2898 GAATTTGTGAGATTAATATTTTAATTAATTAATTTAGATCAATTAATTTTGTCAATGTGAT	2957			
DB	299 GAATTTGTGAGATTAATATTTTAATTAATTAATTTAGATCAATTAATTTTGTCAATGTGAT	358			
QY	2958 TGAATAAATTAATTAATTAATTTCAATTTGAGACAAATTAATGTAATGACCAATTTTATATC	3017			
DB	359 TGAATAAATTAATTAATTTCAATTTGAGACAAATTAATGTAATGACCAATTTTATATC	418			

QY 3018 ATTCTCAAGAGATCAATTAGAGTTTGGTTTAAAAACATGTTGGTATTAATCA 3077  
DB 419 ATTCTCAAGAGATCAATTAGAGTTTGGTTTAAAAACATGTTGGTATTAATCA 478  
QY 3078 TTATTAGTTTCCCAATAATATGAAATA 3106  
DB 479 TTATTAGTTTCCCAATAATATGAAATA 507

RESULT 7  
AY692762 2580 bp DNA linear PLN 11-AUG-2004  
LOCUS Saccharomyces cerevisiae clone FLH158422.01X YOL130W gene, complete cds.  
DEFINITION AY692762.1 GI:51012974  
ACCESSION AY692762  
VERSION Yeast ORF Project.  
KEYWORDS Saccharomyces cerevisiae (baker's yeast)  
SOURCE Saccharomyces cerevisiae  
ORGANISM Saccharomycetes  
Eukaryota; Fungi; Ascomycota; Saccharomycetaceae; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
1 (bases 1 to 2580)  
Marsischky, G., Rolfs, A., Richardson, A., Kane, M., Bagui, M.,  
Taycher, E., Hu, Y., Vanberg, F., Weger, J., Kramer, U., Moreira, D.,  
Kelley, F., Zuo, D., Raphael, J., Hogle, C., Jepson, D., Williamson, J.,  
Camargo, A., Gonzaga, L., Vasconcelos, A.T., Simpson, A., Kolodner, R.,  
Harlow, E. and Labber, J.  
Creation of the YLEX clone resource: cloning of Saccharomyces  
cerevisiae ORFs in the Gateway recombinational cloning system  
Unpublished  
2 (bases 1 to 2580)  
Marsischky, G., Rolfs, A., Richardson, A., Kane, M., Bagui, M.,  
Taycher, E., Hu, Y., Vanberg, F., Weger, J., Kramer, U., Moreira, D.,  
Kelley, F., Zuo, D., Raphael, J., Hogle, C., Jepson, D., Williamson, J.,  
Camargo, A., Gonzaga, L., Vasconcelos, A.T., Simpson, A., Kolodner, R.,  
Harlow, E. and Labber, J.  
Direct Submission  
Submitted (20-JUL-2004) Biological Chemistry and Molecular  
Pharmacology, Harvard Institute of Proteomics, 320 Charles St.,  
Cambridge, MA 02141, USA  
This clone is part of a collection of Saccharomyces cerevisiae  
full-length ORF clones generated by the Harvard Institute of  
Proteomics. Each CDS has been cloned with its native stop-codon.  
The CDS has been directionally cloned using the Gateway cloning  
system into the donor vectors pDONR 201 or pDONR 221. Additional  
sequences in the clone: 'ATCCAGTCAACCC' after the attL1 site and  
before the 'ATG' (from Research Genetics primers) used to amplify  
the ORFs, including a Kozak consensus sequence;  
'ATCCCGGAAATGCCATG' after the stop codon and before the attL2  
site (from the Research Genetics primers used to amplify the ORFs).  
Location/Qualifiers  
1..2580  
/organism="Saccharomyces cerevisiae"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4932"  
/clone="FLH158422.01X"  
/lab\_host="DH5alpha T1 resistant"  
<1..>2580  
/product="YOL130W"  
1..2580  
/codon\_start=1  
/product="YOL130W"  
/protein\_id="AA092781.1"  
/db\_xref="GI:51012975"  
/translation="MSSSSSSSSSSSPNLSRSLANTMYSMKTEHTGLYHRQHPDS  
LVRHQPTLKNKEIAKSTKPSIPKQKATYNNHGVSPSRGMDFFEDGQKMD  
ETVAHQQLASALITSNAPSRSLAHMPOQLYVESNHTPPKVGARVDYTMSSST  
ASSGKSKLASSASAPITKVRKSLVSVLIPHSKSDTSKLAPEKRYSTTSA  
HSINPAVLTKTSQKSDADDTLERKVRNTPASPDVSQASRDSQETREYVCF  
PKPQATRYVNGIDPDELEBYQFANAEKSGTSLAQVNRKYSNVSDIGFTSTST  
TSGSSAAKTYTPRASQTKSKSTNSTBTHKKEDBHEKIRPSLHFGISFGKAYEGE  
ENENIPSDPAYCSYQGTDFQIIPNRPSPSCSSDETVAASDIPSLVSBQTYBLFRG

ORIGIN  
Query Match 11.24; Score 396; DB 8; Length 2580;  
Best Local Similarity 72.5%; Pred. No. 4e-46;  
Matches 513; Conservative 0; Mismatches 195; Indels 0; Gaps 0;  
variation  
942  
/note="compared to SGD sequence"  
/replace="g"  
GEPWWIDSCPTDDMRCIAKAFGHPLEADIDIRMOETREKVELFKSYFVCHTPE  
NDKSSDFLEPILNIVYIVVCSGVLTFFHGPISHCANRRKRRQJLDIVANSDMLCTA  
LIDIDTDFAPVLOSIEYBADIEDSVFAMDMDFAMLDQIGSRKRTWLMKLBSG  
KADYIKMFAKRCODEANGIGPALIOSYENINLQARONASHIKNNSGTPNNYAPTT  
SOPRGIDALYAGDIODHLITMFOULLEYKINIFSSHNTYLAQLOVESFNSNKVTEML  
GKTMIGTMVLPPLNVTIGLFGOMNYKVGENSESIAMWFGIIGLILLVLICMFLASYMI  
KRIDPPATILNEABESGAKSVISFPLPRNKRFRNDRKNINVRAGPSKSVASLPSRIS  
RYD"  
1616 TCTGATGGATTATGAGAGGTCATTACCTGATGATTTCCATTCCTTCATTCGATCA 1675  
DB 1216 TCTTACACAGGTACAGACTTTCAAATTCCTAATGATTTCTTTTCTGTTCCGAATCT 1275  
QY 1676 GAAGAACTATTATGACGCCCGCATATTCATCATTAATGATCACCAGGTCATCTGTTGCA 1735  
DB 1276 GATGAAACCGTTTCATGCTAGATATCCCTTCCTTAGTCTGAGGACAAACATTCAT 1335  
QY 1736 GATTATTATGAAATGTTGAAAGAACTTGTGTTAGATTGTTACTTGTCTACTGATTCG 1795  
DB 1336 GAATTATTCAAGGAGGAGTGACCAACATGCTGTTAGATTGATGTCCTCAACGATGAT 1395  
QY 1796 GAATGAAATAGTTGGCCAAAGCATTTGTTATCATCTTTAAGCTCGGAAGATATTCGA 1855  
DB 1396 GAAATGCTGTCATACGAAAGCGCTTTGGTATTCATTCATGACCTGCGAAGATATTCGA 1455  
QY 1856 ATGCAAGAACTCGTGAAGAAAGTTGAATTTAATTAATTTTGTGTTCCAT 1915  
DB 1456 ATGCAAGAAACGTTGAAGAAAGTTGAATTTAATTTTGTGTTCCAT 1515  
QY 1916 ACTTTGAACTGATTAAGAAATCTGAAGATTATTTAGAACCGATPAATGTTTATTTGTT 1975  
DB 1516 ACCTTTGAAATGATTAAGAAATCTGAAGATTATTTAGAACCGATPAATGTTTATTTGTT 1575  
QY 1976 GTTTTCCATGATGATTAATTAACGTTCCATTTTACCAATTTTCATCCGAAATGTT 2035  
DB 1576 GTCTGTAGTCAGGATGTTTAAACATTTTGAAGCCCAATATTCATGTTGCTTAATGTC 1635  
QY 2036 AGAAGAAAGTTCGTCATTAAGAGATTATGATGATGTTAGTGTGTTAT 2095  
DB 1636 AGAAGCGTGTGAACAGATTACGATGATGATGATGATGATGATGATGATGATGATGAT 1695  
QY 2096 GCCTTAATGATGAATTAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 2155  
DB 1696 GCTTTAATGATGATTAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1755  
QY 2156 GCTGATGATGATGAAGAGCCGTTTTCACGCTGATGATGATGATGATGATGATGATGAT 2215  
DB 1756 GCAAGATGAATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1815  
QY 2216 CAAGAATTTGTTGATCAAGAAGAAAGTCACTTAATGATGATTAATCAAGTAA 2275  
DB 1816 CAAGAATTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1875  
QY 2276 GCTGATGATTAATAATGTTTCTAATAAGATGTCAGAAGAAAGCTAAT 2323  
DB 1876 GCTGATGATTAATAATAATGTTTCTAATAAGATGTCAGAAGAAAGCTAAT 1923

RESULT 8  
SCYOL130W 4202 bp DNA linear PLN 05-AUG-1997  
LOCUS S.cerevisiae chromosome XV reading frame ORF YOL130W.  
DEFINITION SCYOL130W  
ACCESSION Z74872.1 GI:1420016  
KEYWORDS







```

QY 2156 GGTGATGCATTAAGATGCCGTTTCACTGCTAGAGATACATTTAGTAGTATGTA 2215
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2171 GCAAGTGCATGGAAGATTTCTGTCTTCAAGCTCGATGATGATTTGACAACTGTA 2230
QY 2216 CAAAGATTTGGTGAATCAAGAAAGAAAGTCAATGACTTTAATGATTTATCAGGTAAA 2275
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2231 CAAAGATTTGGTGAAGTGAAGGCGTGAAGCAATGACGTTGATGAGACTTCTTAGTGTTAAG 2290
QY 2276 GGTGATGCATTAAGATTTGCTTAAAGATGCTCAAGAAAGACTTAAT 2323
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2291 GGTGATGTTATCAAAATGTTTGCACAAAGATGCAAGATGAAGCTAAT 2338

RESULT 10
CR380951.1
WPCOMMENT
Sequence split into 7 fragments
Fragment Name      Begin      End      LOCUS CR380951 Accession CR380951
CR380951_0         1      110000
CR380951_1        100001      210000
CR380951_2        200001      310000
CR380951_3        300001      410000
CR380951_4        400001      510000
CR380951_5        500001      610000
CR380951_6        600001      687501
Continuation (2 of 7) of CR380951 from base 100001 (CR380951 Candida glabrata strain CBS

Query Match      10.9%; Score 385; DB 8; Length 110000;
Best Local Similarity 72.8%; Pred. No. 4,8e-45;
Matches 496; Conservative 0; Mismatches 185; Indels 0; Gaps 0;

QY 1633 CCTGATAGTTTTCATCTTTCATCTGATATGATGAAAGAAACATATTCATGCCCCCATAT 1702
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 55814 CCTGACAAATTTCTTCTTCCATCAGAAATGAGAAAGAAACATATTCACAGATATA 55873
QY 1703 CCATCATTTAGTATCAACAGATCAATCTGTTGCGAGATTTATTTAGAAATGATGAAGAACT 1762
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 55874 GCAACTTAGTACACACAGGAGATCATCTTTGACCTTTCAAGGGGCGGAGAGCCACAG 55933
QY 1763 TGGTGGTAGATTTGATCTTGTCTTCTACTGATTCGAAATGAAATGTTGGCCCAAGCATTT 1822
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 55934 TGGTGGTAGATTTGATCTTGTCTTCTACTGATTCGAAATGAAATGTTGGCCCAAGCATTT 55993
QY 1823 GGTATTCATCTTTTAACTGCTGAAGATATTCGAATGCAAGAAACCTGCTGAAAAAGTTGAA 1882
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 55994 GGCATTCATCTTGTGACCTGCTGAAGATATTTAGAAATGCAAGAAACCTGCTGAAGAGTGAG 56053
QY 1883 TTAATTTAAAGTTATTTATTTGTTTCCATACCTTTTGAAGCTGATTAAGATCTGAA 1942
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 56054 TTAATTTAAAGTTATTTATTTGTTTCCATACCTTTTGAAGATCTGAA 56113
QY 1943 GATTAATTTAGAACCGATAAATGTTATATTTGTTTCCATGATGCTATATTTAAAGCTTTC 2002
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 56114 GATTTCCTAGAGCCCAATGATATATGAGTTGATTTAGATCAAGGTTTGTTCACATTT 56173
QY 2003 CATTTTTCACCAATTTCTATCAGCAAAATGTTAGAAAGAGTTGCTGTAATTTAGAGAT 2062
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 56174 CACTTTTGACCAATTTCAATTTGCTATATTTAGAAAGCGTGTGAGACGCTGCGAGAT 56233
QY 2063 TATGTCAGATTTAGAGCTGATTTGTTTATGCTTAAATGATGAAGAAATTCACCAATGAT 2122
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 56234 TATGTCAGATTTAGAGCTGATTTGTTTATGCTTAAATGATGAAGAAATTCACCAATGAT 56293
QY 2123 TTTGGCCCCGATTCATGGAATTTGAATTTGAAGCTGATGCAATTTGAAGATCCGTTTTC 2182
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 56294 TTTGGCTCAGATTATATGATGATGATGAGCTGAAGAGGAGCTCAATTTGAAGACCTCGGTTT 56353
QY 2183 ACTGCTAGAGATACGATTTTATGTAATTTGTTTCAAAAGATTTGCTGATCAAGAAAGAAA 2242
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 56354 ATGGGAAGAGACTTTGATTTTCTGCGATTTCAAAAGATCGGTGAAGAGTTAGGGCGCAAA 56413
QY 2243 GTCATGATCTTTAATGATATTTATTCAGGTAAAGCTGATGTCATTTAAATGTTTGTCTAAA 2302

```

```

Db 56414 ACGATGACCTTTAATGATTAAGTTACTTAAGTGGCAAGACGATGTCATCAAAATGTTCCCAAG 56473
QY 2303 AGATGCTCAAGAAAGACTTAAT 2323
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 56474 AGATGCTCAAGAAAGACTTAAT 56494

RESULT 11
YSCCHRVIN/C
LOCUS      270148 bp      DNA      linear      PLN 29-MAY-2002
DEFINITION Saccharomyces cerevisiae chromosome VI complete DNA sequence.
ACCESSION   D50617 D31600 D44594 D44595 D44596 D44597 D44598 D44599 D44600
            D44601 D44602 D44603 D44604 D44605 D44606
VERSION      D50617.1 GI:2804269
KEYWORDS     sulfite reductase flavoprotein; mitochondrial DNA-directed RNA
            polymerase; 12KD heat shock protein; 96KD nucleoporine-interacting
            component; DNA polymerase epsilon suppressor 4; FBI1 protein;
            GTP-binding protein YPT1; MDJ1 protein precursor; Muts protein
            homolog 4; actin; bZIP protein binding to CRE motif;
            calcium/calmodulin-dependent protein kinase type I; cell division
            control protein 14; cell division control protein 4; cell division
            control protein CDC26; chromosome segregation protein SMC1;
            chromosome segregation protein SMC2p; cytoplasmic phenylalanyl-tRNA
            synthetase beta chain; depressed growth-rate protein;
            dihydroliposamide dehydrogenase precursor; glycogen synthase isoform
            1; hexokinase A; histidinol phosphatase; mitochondrial ribosomal
            protein YMR-31 precursor; nuclear integrity protein 1; pheromone
            alpha factor receptor; phosphate system positive regulatory
            protein; phosphomannomutase; proteosome component PRE4; RNA
            helicase; Ras-related protein; transposon TY1-17 154.0KD
            hypothetical protein; transposon TY1-17 49.8KD hypothetical
            protein; tubulin beta chain; ubiquitinol-cytochrome C reductase 17KD
            protein.
SOURCE      Saccharomyces cerevisiae (baker's yeast)
ORGANISM   Saccharomyces cerevisiae
REFERENCE 1 Murakami,Y., Naicou,M., Hagiwara,H., Shibata,T., Ozawa,M.,
            Sasanuma,S., Sasanuma,M., Tsuchiya,Y., Soeda,E., Yokoyama,K.,
            Yamazaki,M., Tashiro,H. and Eki,T.
            Analysis of the nucleotide sequence of chromosome VI from
            Saccharomyces cerevisiae
            Nat. Genet. 10 (3), 261-268 (1995)
            95400292
            MEDLINE 7670463
            PUBMED 2 (bases 1 to 270148)
            REFERENCE Murakami,Y.
            Direct Submission
            TITLE Submitted (23-MAY-1995) Yasufumi Murakami, Tsukuba Life Science
            Center, RIKEN, Division of Human Genome Research, 3-1-1 Koyadai,
            Tsukuba, Ibaraki 305, Japan (E-mail:yasufumi@rclcs.riken.go.jp,
            Tel:81-298-36-9059, Fax:81-298-36-9137)
            On or before May 14, 2001 this sequence version replaced gi:836814,
            gi:2804274, gi:1100783, gi:871938, gi:871933, gi:1122225,
            gi:870817, gi:1122230, gi:893419, gi:871957, gi:870821, gi:1122231,
            gi:83685.
            D31600:Submitted (25-MAY-1994)
            D44594-D44606:Submitted (28-Dec-1994)
            Sequence updated (16-Jan-1998)
            Sequence updated (28-Dec-1995) by: Yasufumi Murakami
            Sequence data v3.0.
            We have corrected 2 sequence errors within some CDSs. Therefore, we
            have revised 3 CDSs (YFL042C,43C,31W). YFL043C were connected to
            YFL042C and eliminated. The length of YFL031W (HAC1) became more
            longer than previous version.
FEATURES
            source
            location/Qualifiers
            1..270148
            /organism="Saccharomyces cerevisiae"
            /mol_type="genomic DNA"
            /strain="AB972"
            /db_xref="taxon:4932"

```



```

misc_feature
1..4685
/note="y' element"
CDS
33..535
/note="YFL068W"
/codon_start=1
/protein_id="BAA09173.1"
/db_xref="GI:836687"
/translation="MMPAKLQDVLRTIQSSARHGQTQLKNSNPLERPHKDRIVYECLEP
FPFALFLVGVQVLOHLCIRFQVAPYPIQLFDPSPRHAENLAPLASCRIOYNCF
SSSSNGQVPSITSLYLRVDLSPFYAKKQIPIRVEMIMLDVQVFEVLVYSQSLHS
"
CDS
836..1363
/note="YFL067W"
/codon_start=1
/protein_id="BAA09174.1"
/db_xref="GI:836688"
/translation="MESTITLAIPIGVLIGTVAGAGSSSIPDVAGSGSRSPDV
DAGSGSRISAGVITSGSSTSPDVAGSGSSTSPDVAGSGSSISAGVSRIGTIGS
RIGTGIGTIGSISTSGRISIPDVGSNGRISITGVSTGISTMARVAVLITAAI
LSAPYTAIALLEARR"
complement(1437..2615)
/note="YFL066C"
/codon_start=1
/protein_id="BAA09175.1"
/db_xref="GI:836689"
/translation="MADTPSVAVQAPGYGKTELEHPLPLIALASKGVYVSFLFVPY
TYLANCMRLGRGCLNAPVRNPIEBGYDGVTLVYGIYDLASTFTTRIAMEN
IYECPTPTNNTVGLIYDEPHNFEYVROGFGDITLDEPAEKAFISGTAPEA
VADALORIGTGLAKKMDINELRSDELRSGLSYPTPMNLIKESKSPYGLVHK
IKTKYTSQPEBALIKLILAFIPIBPSKAIYVASTNVEBELACSWRKTRIVYVINGKL
GAARVSRTRKEFVTDGSMQVLIIGTKLVTEGIDIKOLMNYTMDNRLNIIELIQGRL
RGGGLCYLSRNSMAARNRKELPLPIBEGCITBQVREPYGLSKKGGQHVGCVC
"
CDS
complement(1030..3338)
/note="YFL065C"
/codon_start=1
/protein_id="BAA09176.1"
/db_xref="GI:836690"
/translation="MRTPTDPVSGAPIYVSLQKSTIRKGYNLAPHMLLHVDELSI
PSAYQASLPEKKVTERLKRDLCPRKPIEIKYFQIICNDMMNKKDRIGDVLVCCPS
"
complement(1322..3846)
/note="YFL064C"
/codon_start=1
/protein_id="BAA09177.1"
/db_xref="GI:836691"
/translation="MKVSDRRKPEKRNPDRESALNNKNDLVHCPSITLFSIPTEV
STYEDBSGLIKVKTGAMNRKSPFKIYISVWGVQVQFLFVEDDEDPQGGPI
PKRYLIPKINIMVYTLFQVHTLKNRKDYDLSFLNRYNELSPPCPTLSRNS
ECCAROLYDAVFH"
4685..4823
/note="left telomere"
4823..5000
/note="X element"
5066..5521
/note="YFL063W"
/codon_start=1
/protein_id="BAA09178.1"
/db_xref="GI:836692"
/translation="MIVNNAHILTPPIPHAPSVLITTCILIRHTHTDATTVTITST
YPTTFHSHMAHLSHEVYCTHIIMGTCLSGYVPEFTNHSYEPHFNYILFRPKY
CTNALTNTVPIPHRILTYQFIYVANTIKSPKPLSKPKHNILSFNNNT"
6426..7565
/note="YFL062W"
/codon_start=1
/protein_id="BAA09179.1"
/db_xref="GI:836693"
/translation="MKENELKNKSKSVYVLSFKQLBSQKIVLPQDLFRSSFTWFCYEIY
KSLAPILMLMLPLSVWKLNNMTIPLWVSLVLFMGVPLVLYFRLSRGSRISKO
LFOCKEITKSTPSSDPHDMEVAAALNGLYENKRNARYFFPNAMGCOBARFTTL
EFSPLKDEAAKYSFKDSVPYTERALGYPRRVEKQMLTPNSSEKMSVYGJEDAKLP
KSAVRKLTWLPKRISNIFMLIPFLNLCCTIVSRQMCLLNLTLYGLVFWQGFQ

```

```

CDS
NIRVLIMSMEHKQFLSTIINEBESGANGDEILARKNRYLPEKKVWKNIEFPDGIID
CEMFHSHPFTRVYLSAKKSMRSLSNVLEWLPYIKKAQSLCSBESLA"
9545..10222
/note="YFL061W"
/codon_start=1
/protein_id="BAA09180.1"
/db_xref="GI:836694"
/translation="MSQYGFVRVREVEKAIPIVNAAPRAVVPVPSSEFARLYREYA
AKELTYAVLNHSLNLFQYSAIIRDOFPADLDQEVLYVTCLLHDIAITDNMRATM
SFEYVGGIILSLBELVFNATGQADYADATEAIIRHODLTGVTGLIQTGLIQTATTL
NVSNTDLIHDIYSAINEQFPRLHMLSCEFATVVDIENSRRKPMWHTSLSLGDPSKVI
CNTRGYN"
complement(10301..10969)
/note="YFL060C"
/codon_start=1
/protein_id="BAA09181.1"
/db_xref="GI:836695"
/translation="MTVIVIGLALQGAFFIEHVRHVEKCIYENRDFYEKLSVMTVKDK
NQLAQCDALITPGESTAMSLIARNGPYDLYAFVNPSPKVTMGTCAGLIVISQULS
NEALVLTMLNLRKVRKNAFROAGSTRICDSNTIPHCNDPATPIRAPVYEVLY
DPEHVQVLYKLDGKNGQQLIVAKOKNNIATSPHELAENDIRPHDWIRFELVK
NYSK"
11363..12259
/note="YFL059W"
/codon_start=1
/protein_id="BAA09182.1"
/db_xref="GI:836696"
/translation="MSBFRKVTGLAQLMKGVINDVYTPQQAIIAERAGACAVALLER
IPAMRSGQVCRMSDPMIKEIMEAVSIPIVMAKVRIGHFEAQIILELQVYIDBS
VLTPADTWHIEKNPFKVPFVGAKDGEALRINEGAMARTGEGAGTGVSAVKH
ITRKADICQYKENTLBSDEPAKATELRVPVDLTKTILSGKLPIVNFAGAVATPA
DALIMQDCBQVYVGGISIPKSPBKLCAIYVATTHYNDPALQVSSLDGLMGC
ISIOSINBAGKNAARLSEIGW"
12929..13951
/note="YFL058W"
/codon_start=1
/protein_id="BAA09183.1"
/db_xref="GI:836697"
/translation="MSTDKITFLNMQPTPYHIDIPLAQTKYREBQGLDMAILBPTN
PSDVTIELSGSKVMGKAMHTITLAAKRGFPVTSVSLDEPRTGVLYLKGSGITGD
FQSLKGGKIGVGEFGKQIDELTKHYGMKEDYTVASCNVAKYIIEGKIDAGIGI
KMOQVLEEVLYAKORGASPADKMLRIDKLCAGCCGCTVLYCNDPEFLKNEPKVR
KFLAKIKATHTYVADPYKMKRYIDPRQALNCLSGCYKQYCAVYESSLYNHRDW
KKTGYGRALILPPDYVSNYTNELTSLPPEEVSVDPLEARLMAIHQEKROGTFR
RLALPA"
complement(14305..14763)
/note="YFL057C"
/codon_start=1
/protein_id="BAA09184.1"
/db_xref="GI:836698"
/translation="MARHFGMALAPMDVMGGGFQSKKAMEERRKNGEIRSPVQASB
QTDABIKISALAKVAEBHGTESYTAIAIAYRSKAKNVFLVGGRIEHLKONIEAL
SIKLTPBOIKLIESIIPDVGFPTNFIGDDPAVTKKASILTAMSAQISFD"
complement(14793..15431)
/note="YFL056C"
/codon_start=1
/protein_id="BAA09185.1"
/db_xref="GI:836699"
/translation="MADLFAPEPSTELGRLRYLSKAGIRVSPLLIGMSIGDAMS
ELIGMSKERAPFLIDAFYAGNPIIDTANNYONBOSAWIGEMVWSKLDQIVYAT
KFTDYKRYDVGGSKSNYCGNHRKSLHVSRLRCLQTDWIDILVYHMDVYSSIE
EVMQSLHLVLOQSSSIVLCIRLPLGLRQITTLNLMVFKLASIKVNGTC"
17004..18680
/note="YFL055W"

```

CDS

Query Match 10 9%; Score 383; DB 8; Length 270148;  
 Best Local Similarity 59.8%; Pred. No. 7e-45; Indels 165; Gaps 3;  
 Matches 803; Conservative 0; Mismatches 375;

1625 ATTATGAGGCTTCTTACCGATAGATTTTCACCTTTCACATTCGATCGAAGAAACT 1684  
 |||||  
 34627 AATACGAATATCAAAATTCGATGATTTTCGTTCTTCGCGTCTGAATCGATGAAACT 34568  
 |||||

QY 1685 ATTCATGCCCCGATATTCATCATTTAGTATCAACGAGCTCAATCTGTTGAGATTATTT 1744  
 DB 34567 GTGATGCTAGATATATCCCTTCTTTGATATCCGAGGAGACAGCTTTATGAGATTATTC 34508  
 QY 1745 AGAATGCTGAGAAATCTTGCTGTTAGATGTTACTTGTCTTACTGATTCGGAATGAAA 1804  
 DB 34507 AAGGAGGCGATCCAACTGGGGTGTGATGTCAGTTGTCCAACTGATGAGAAATGCGT 34448  
 QY 1805 ATGTTGCGCAAGCAATTTGATTCATCTTAACTGCGAAGATATTCGATGCAAGA 1864  
 DB 34447 TGCATAGCAAAACATTTGAAATCATCTTTGACCGGAGAAATATCAGATGCAAGA 34388  
 QY 1865 ACTGCGTAAAAAGTGAATTTATTTAAAGTTATTTTGTGTTTCCATATCTTTGAA 1924  
 DB 34387 ACGGCTGAAAAAGTGAAGCTTTTCAATCCATATCTTTGTCTGTTCCATATCTTTGAA 34328  
 QY 1925 GCTGATTAAGAACTGGAATTTATTTAGAACCGATTAATGTTTATTTGTTTCCAT 1984  
 DB 34327 AATGATTAAGAACTGGAATTTATCTTGAACCTATTAATGTTTATCTGTTTATAG 34268  
 QY 1985 GATGATATTTAAGCTTTCATTTTCCACCAATTTTCCATCCAGCAAAATTTAGAGAA 2044  
 DB 34267 TCAGGTGCTTGAATCTTTTCAATTTTGAACCAATATCACTGTGCAATGTAAGAAAGCT 34208  
 QY 2045 GTTGTGCTCAATGAGAGATTTATGTCATGTTAGTGTGCTGATTTGTTATGCTTAATC 2104  
 DB 34207 GTAGACATTTATGCGATCTATGTCAGTGTCAATTCAGATGTTATGTTATGCTTAATC 34148  
 QY 2105 GATGAAATTTACGATGTTTGTGCCCCGCTGATTCATGAAATTTGAAATGAACTGATGCC 2164  
 DB 34147 GATGATATTCACGATGATGTTTGTGCTCGGTTATTCATCATGATGATGAACTGATTCG 34088  
 QY 2165 ATGGAAGATGCGCTTTCATCTGCTAGAGATCTGATTTTATGATGATGTTTAAAGATTT 2224  
 DB 34087 ATGGAATGATGATGTTTATGAGACCGGTATATGATGATTTGACCACTGTTTAAAGATTT 34028  
 QY 2225 GGTGAATCAAGAAAGAAATGTCATGCTTTATGATGATTTATTCAGTAAAGCTGATGTC 2284  
 DB 34027 GGTGAATGAGCGCTTATGAGATGACGCTGATGATGATCTTTGAGCGGTAAAGCAATGTT 33968  
 QY 2285 ATTAATATGTTTGTCTTAAAGATGTCAGAAAGAAAGTAACTCTTCTGTTATATCA 2344  
 DB 33967 ATCAAAATGTTGCGCAAAAGATGTCAGAAAGTAACTAAATGATTTGGGCACTGCTTAA 33908  
 QY 2345 GGTCAATTTACTTTCACACACACACACAGGCCCAACCAACCACTTAACTTATTT 2404  
 DB 33907 TCTCAAACTTAACTGATGAACTTACAA 33881  
 QY 2405 ATTACTTCAACCAATTTATTCACCTTGAATCTTAATGTTTGAATCTTCAACTGATGGA 2464  
 DB 33880 33881  
 QY 2465 GGAGTAGAGTAGAGAAATTAATTTTGTGCTCCAACTCAACTGAAATTAATTAATCT 2524  
 DB 33880 33881  
 QY 2525 AATTAATACTACTGCTTCACTTCACTCACTCACTCACTCACTCACTCACTCACTCACT 2584  
 DB 33847 AATAATCAAAATAGCTTATCAATAGCTATATGCAAAACAACTCACTCACTCACT 33800  
 QY 2585 AACAATCTTTCCCATCCCGATGCAAGCTGCAAGCTGATTTGATTTATTTAGGT 2644  
 DB 33799 33770  
 QY 2645 GATATTTCAAGATCATATATATCACTGTTTCAAAATTTATTTAGCTTATGCAAAATTTTTC 2704  
 DB 33769 GATATTTCAAGATCATTTGTTGATGATGTTTCAAAATTTATTTAGCTTATGCAAAATTTTTC 33710  
 QY 2705 AGTGTGCTCACTTCAATTTATTTAGCTCACTTCACTTCACTTCACTTCACTTCACTTAT 2764  
 DB 33709 TCTATATCACTGCAAAATCTACTGCGCAATTTACAAAGTGAATCACTTTAACTCTATATAT 33650  
 QY 2765 AAATACCGAAATGTTTCTAAATTTACTTTGATTTGGGCAATGTTAGTCCATTAAT 2824

DB 33649 AAAGTTACAGAAATGTTGGGTAAGTTACTATGTTAGGTATCAATGTTATCTTCAATTAAT 33590  
 QY 2825 TTATGACAGGACCTTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2884  
 DB 33569 GTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 33533  
 QY 2885 GGTGCTTTTTCGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2944  
 DB 33532 GGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 33473  
 QY 2945 GCTCAATGCTGCTGCAAAATTT 2967  
 DB 33472 GCTAGCTATGATGATCAAAATTT 33450

RESULT 12  
 CR382125\_06/c  
 WPCOMMENT  
 Sequence split into 23 fragments LOCUS CR382125 Accession CR382125  
 Fragment Name Begin End  
 CR382125\_00 1 110000  
 CR382125\_01 100001 210000  
 CR382125\_02 200001 310000  
 CR382125\_03 300001 410000  
 CR382125\_04 400001 510000  
 CR382125\_05 500001 610000  
 CR382125\_06 600001 710000  
 CR382125\_07 700001 810000  
 CR382125\_08 800001 910000  
 CR382125\_09 900001 1010000  
 CR382125\_10 1000001 1110000  
 CR382125\_11 1100001 1210000  
 CR382125\_12 1200001 1310000  
 CR382125\_13 1300001 1410000  
 CR382125\_14 1400001 1510000  
 CR382125\_15 1500001 1610000  
 CR382125\_16 1600001 1710000  
 CR382125\_17 1700001 1810000  
 CR382125\_18 1800001 1910000  
 CR382125\_19 1900001 2010000  
 CR382125\_20 2000001 2110000  
 CR382125\_21 2100001 2210000  
 CR382125\_22 2200001 234072

Continuation (7 of 23) of CR382125 from base 600001 (CR382125 Kluveromyces lactic strair  
 Query Match 10.7%; Score 376; DB 8; Length 110000;  
 Best Local Similarity 70.5%; Pred. No. 8.7e-44;  
 Matches 502; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

QY 1612 AATTTCTGATGGAGTATATGAGAGTTCATTAACCTGATGATGATTTTCACTTTCCATTGGA 1671  
 DB 59964 AATACCAAGAAAGAACTTCACTTATGATGACCACTAATGATTTGCTTATTTAGTTGGA 59905  
 QY 1672 ATCAGAAAGAACTATTCATGAGCCCGATATTCATCATTTAGTATCAGAGCTCAATCTGT 1731  
 DB 59904 CGTATCTGAAACTTTGATGCTCCAGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 59845  
 QY 1732 TCGAATTTATTTAGAAATGTTGGAAGAACTTGTGTTAGATTTGCTTACTGCTGCTGCTG 1791  
 DB 59844 TAGACAAATTTATTCATGATGATGCTCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 59785  
 QY 1792 TTGGAATTTGAAATTTGTTGCTGCAAGCACTTTGCTATTTGATCTTTTAACTGCTGGAATAT 1851  
 DB 59784 TGAGAAATTTGAGTGTGCTTACCAAACTTTGGAATTTATCTTGAACATTCAGAAAGATAT 59725  
 QY 1852 TCGAATTTGAAAGAACTGCTGCAAAAGTTGAAATTTATTTAAAGTTATTTTGTGTTGTTT 1911  
 DB 59724 CAGATTTGAAAGAACTGCTGCAAAAGTTGAAATTTTCAAAACATTTATTTTCACTGCTGTT 59665  
 QY 1912 CCATATCTTTGAAAGCTGATTAAGAAATCTGGAAGATTTTATGAAACCGATTAATGTTTATAT 1971  
 DB 59664 CCATATCTTTGAAAGCTGATTAAGAAATCTGGAAGATTTTATGAAACCGATTAATGTTTATAT 59605

```

QY 1972 TGTGTTTCCATGATGTAATTAACGTTCCATTTTCCACCAATTTCTCATCCAGCAA 2031
DB 59604 GGTGTTTTTGGCAAGGATTTCTAACTTTCACTTAAACCAATCTCAATTTGGCCAA 59545
QY 2032 TGTGGAAGAAGATTTGCTCAATTTGAGAGATTAATGTCATGTTAGTGTGATTTGTTATG 2091
DB 59544 TGTCAAGAAGACGTGTGAGCAACTAGAGATTAATGTGGCGTAACTCAAGATTTGATTTG 59485
QY 2092 TTATGCTTAATCGATGAAATTAACGATGTTTGGCCCGGATTCATGAAATTTGAATA 2151
DB 59484 TTACGCTTAGAGATGATATCAAGATGTTTGGACCTGTATTTCAATCATGATTA 59425
QY 2152 TGAAGCTGATGCTGATGAAAGATGCGTTTCACTGCTAGAGATGATTTTATGATAT 2211
DB 59424 TGAAGCAATGCAATTAAGAAATTCAGTGTATTATGACAGAGACATGAATTTGCTGATAT 59365
QY 2212 GTTACAAAGAAATTTGCTGATGCAAGAAAGAAAGTCAATGATTTAATGATTTATCATG 2271
DB 59364 GCTTCAAGAAATTTGGTGAAGAAAGAAAGAAACAAATGACATTAATGATTAATCATG 59305
QY 2272 TAAAGCTGATGCTTAAATGTTTGTCTAAAGATGTCAGAAAGAAAGTAAAT 2323
DB 59304 TAAAGCAATGTTATCAAAATGTTTGTCCAAAGATGTCAGAGCATATCAT 59253

```

RESULT 13  
AE016817.03

Sequence split into 15 fragments LOCUS AE016817 Accession AE016817

```

Fragment Name Begin End
AE016817_00 1 110000
AE016817_01 100001 210000
AE016817_02 200001 310000
AE016817_03 300001 410000
AE016817_04 400001 510000
AE016817_05 500001 610000
AE016817_06 600001 710000
AE016817_07 700001 810000
AE016817_08 800001 910000
AE016817_09 900001 1010000
AE016817_10 1000001 1110000
AE016817_11 1100001 1210000
AE016817_12 1200001 1310000
AE016817_13 1300001 1410000
AE016817_14 1400001 146886
Continuation (4 of 15) of AE016817 from base 300001 (AE016817 Eremothecium gossypii ATCC

```

Query Match 10.6%; Score 373.6; DB 8; Length 110000;  
Best Local Similarity 71.6%; Pred. No. 1.9e-43;  
Matches 490; Conservative 0; Mismatches 194; Indels 0; Gaps 0;

```

QY 1639 ATTACCTGATGATTTTTCACCTTTTCATTTCTGAATTCAGAAAGAACTATTCATGCCCCCA 1698
DB 94322 AATACCTGATGATTTTTCCTTTCTGTTCCAGAGAGAGAGACATGTCATGCCCTGA 94381
QY 1699 TATTCATCATTAAGTATCAACAGATCATCTGTTGAGATTTATTTAGAAATGTTGAGA 1758
DB 94382 TATTCATCTCTTTGTCAGAGATGAGAAACCTTCCTGTAATTTTGAAGATGTTGCCGT 94441
QY 1759 AACTGTGATGATGATTTGATGATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1818
DB 94442 TACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 94501
QY 1819 ATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1878
DB 94502 TTTTGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 94561
QY 1879 TGAATTAATTAAGTATTAATTTTGTGTTTCCATCTTTTGAAGTGAATTAAGTAT 1938
DB 94562 TGAATTTTAAATCACTACTCTTTGTGTGTTTCTACTTTTGAAGTGAATTCGGAATC 94621
QY 1939 TGAAGATTAATTAAGAACGATTAATGTTTATGTTGTTTCCATGATGATTAATTAAC 1998

```

```

DB 94622 TGCAGACTTTTGGAGCCCATCAATGATATGTTGTTCTTCAGAAATGATGTTTGCAC 94681
QY 1999 GTTCCATTTTTCACCAATTTTCTCATCCAGCAATGTTAGAAAGAGTCTGCAATTTAG 2058
DB 94682 CTTCACCTTTAGCCAGATGTTCTCATCTCCCAATGTCAGAGCCGTGTTGCCAGTTTGA 94741
QY 2059 AGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2118
DB 94742 GGACTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 94801
QY 2119 TGTGTTTGGCCCGGATTCATGAAATGATTAAGAGCTGATCCATGTAAGATGCCGT 2178
DB 94802 TAGTGTCTCAGTAATTAATCAACTATTAATGAAATGCAAGCCATGTCATGAAAGTTG 94861
QY 2179 TTTCACTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2238
DB 94862 TTTCACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 94921
QY 2239 AAAAGTCATGCTTTAATGATTAATGATTAATGATGATGATGATGATGATGATGATGAT 2298
DB 94922 GAAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 94981
QY 2299 TAAAGATGTCAGAAAGAGCTAA 2322
DB 94982 TAAAGCTGTCAGATGAGCAAA 95005

```

RESULT 14

LOCUS CO446161 684 bp DNA linear PAT 30-JAN-2004  
DEFINITION Sequence 11921 from Patent WO0192523.  
ACCESSION CO446161  
VERSION CO446161.1 GI:41417311  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
1 Shimkets R.A. and Leach M.D.  
Human polynucleotides and polypeptides encoded thereby  
Patent: WO 0192523-A 11921 06-DEC-2001;  
Curagen Corporation (US)  
Location/Qualifiers

FEATURES  
source

1..684  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 7.5%; Score 263.8; DB 6; Length 684;  
Best Local Similarity 70.0%; Pred. No. 1.8e-27;  
Matches 355; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

```

QY 1865 ACTCGTGAAGAAAGTTGAATTAATTAAGTATTAATTTGTTGTTTCCATACCTTTGAA 1924
DB 1 ACGCGTGAAGAAAGTTGAGACTTTTCAATCTATTTCTGTTGCTGTTCCATACATTTGAA 60
QY 1925 GGTGATTAAGATCTGAAGATTTATTTAGAACCGAATTAATGTTTATTTGTTTCCAT 1984
DB 61 AATGATTAAGATCTGAAGATTTATTTGAACTTAATTAATGTTTATTTGATGTTTAAAG 120
QY 1985 GATGATTAATTAAGTTCATTTTTCACCAATTTTCTCATCCAGCAATGTTGAGAAGA 2044
DB 121 TCAAGTGTCTGATCTTTTCACTTTTGAACCAATATACACTGCGCAATGTAAGAAAGCT 180
QY 2045 GTTGTCAATTAAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2104
DB 181 GTAAGACATTAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 2105 GATGAATTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2164

```

Db 241 GATGATATACAGATAGTTTGGCTCCGGTATTATTCATTCATCGAGTATGACCTGATTCG 300

Qy 2165 ATTTGAAGTCCGCTTTTACCTGCTAGATCTAGTATTTAGTATGTTTACAAAGATT 2224

Db 301 ATTTGATGATTCAGTGTATTATGACCCGTGATATGATTTTGCAGCCATGTTTACAAAGATT 360

Qy 2225 GGTGAATCAAGAAAGAAAGTCACTTAAATGATTTATTTATCAGTAAAGCTGATGC 2284

Db 361 GGTGAATGAGCCGCTTACGATGATGACGCTGATGACACTCTTACGGGTAAACAGATGTT 420

Qy 2285 ATTTAAATGTTTGTCTTAAAGATGTCAGAAAGAGCTAATTTCTTCTGCTATTATTCAA 2344

Db 421 ATCAAAATGTTGCGCAAAAGATGTCAGATGAACTAATGATGTTGCGCAGTCTTAA 480

Qy 2345 CGTCAATATTAATTTACACACACAA 2371

Db 481 TCTCAACTAATACATGCTGTAACCTTACA 507

RESULT 15

SPBC27B12/ 35030 bp DNA linear PLN 23-JUN-2003

LOCUS SPBC27B12

DEFINITION S.pombe chromosome II cosmid c27B12.

ACCESSION AL021766

VERSION AL021766.1 GI:2853106

KEYWORDS

SOURCE Schizosaccharomyces pombe (fission yeast)

ORGANISM Schizosaccharomyces pombe

REFERENCE 1 (bases 1 to 35030)

AUTHORS Wood, V., Galliam, R., Rajandream, M.A., Lyne, M., Lyne, R., Stewart, A., Soutos, J., Peat, N., Hayles, J., Baker, S., Basham, D., Bowman, S., Brooks, K., Brown, D., Brown, S., Chillingworth, T., Churcher, C., Collins, M., Connor, R., Cronin, A., Davis, P., Felwell, T., Fraser, A., Gentles, S., Goble, A., Hamlin, N., Harris, D., Hidalgo, J., Hodgson, G., Holroyd, S., Hornby, T., Howarth, S., Huckle, E., J., Hunt, S., Jagsels, K., James, K., Jones, L., Jones, M., Leach, S., McDonald, S., McLean, J., Mooney, P., Moule, S., Mungall, K., Murphy, L., Niblett, D., Odeli, C., Oliver, K., O'Neill, S., Pearson, D., Quail, M.A., Rabinowitsch, E., Rutherford, K., Rutter, S., Saunders, D., Seeger, K., Sharp, S., Skelton, J., Simmonds, M., Squires, R., Squares, S., Stevens, K., Taylor, K., Taylor, R.G., Tivey, A., Walsh, S., Warren, T., Whitehead, S., Woodward, J., Volckaert, G., Aert, R., Robben, J., Gymnopoulos, B., Welljens, I., Vanstreels, E., Rieger, M., Schaefer, M., Muller-Auer, S., Gabel, C., Fuchs, M., Dusterhoft, A., Fritze, C., Holzer, E., Moestl, D., Hilbert, H., Borzym, K., Langer, I., Beck, A., Lehrach, H., Reinhardt, R., Pohl, T.M., Eger, P., Zimmermann, W., Medler, H., Wambut, R., Purnelle, B., Goffeau, A., Cadieu, F., Dreano, S., Gloux, S., Lelaure, V., Mottier, S., Gilbert, F., Aves, S.J., Xiang, Z., Hunt, G., Moore, K., Hurst, S.M., Lucas, M., Rochet, M., Galliard, C., Tallada, V.A., Garzon, A., Thode, G., Daga, R.R., Cruzado, L., Jimenez, J., Sanchez, S., Amstrong, J., Forsberg, S.L., Cerutti, L., Revuelta, J.L., Moreno, S., del Rey, F., Benito, J., Dominguez, A., Lowe, T., McCombie, W.R., Paulsen, I., Potashkin, J., Shpakovski, G.V., Ussery, D., Barrell, B.G. and Nurse, P.

TITLE The genome sequence of Schizosaccharomyces pombe

JOURNAL Nature 415 (6674), 871-880 (2002)

MEDLINE 21848401

PUBMED 11859360

REFERENCE 2 (bases 1 to 35030)

AUTHORS Wood, V., Rajandream, M.A., Barrell, B.G., Lauber, J., Hilbert, H. and Dusterhoft, A.

TITLE Direct Submission

JOURNAL Submitted (05-DEC-1997) European Schizosaccharomyces genome sequencing project, Sanger Institute, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: pombe@sanger.ac.uk and QIAGEN GmbH, Max-Volmer-Str 4, D-40724 Hilden, Germany

COMMENT Notes:

Details of S. pombe sequencing at the Sanger Institute are available on the World Wide Web.

## FEATURES

source

1. 35030

/organism="Schizosaccharomyces pombe"

/mol\_type="genomic DNA"

/strain="972h"

/db\_xref="taxon:4896"

/chromosome="II"

/map="ITL"

/clone="cosmid c27B12"

1. 2559

/note="nominal overlap with cosmid SPBC30B4, EM:AL031262 S.pombe chromosome 2"

/complement(join(1..81,146..886))

/gene="SPBC27B12.01c"

/note="synonym: SPBC30B4.09c"

/complement(join(1..81,146..886))

/gene="SPBC27B12.01c"

/codon\_start=1

/protein\_id="CA16896.1"

/db\_xref="GI:2853107"

/translation="MIHLPGQSFQGLIVGQLTLAIYVLPFLFQSPPIPKSVANS PKQNGEPDPTSTPLNNKKRYKPLTLEPILNLYDVNHEPSPIDFVNLIA OALIQRYDACSNDVALRKLETYLNKGQDSMDVHIYVRLDGLDGPVPSGVLP HOHNSQIRAEMLVSLTNINCTYDTKLINPRKAPRTLSIVRICKPYGXSLF TTLIVYSYLDSPSPGLKHQVLRITVIMYFSPSGAGOPAYMNLSPDPNFI"

/complement(82..97)

/gene="SPBC27B12.01c"

/note="ctacacttgcctcg, splice branch and acceptor"

/complement(140..145)

/gene="SPBC27B12.01c"

/note="gtatic, splice donor sequence"

1053..1064

/note="12 Ts in SPBC27B12 and SPB4539, but only 11 in SPAC30B4, we have altered c30B4 to agree with the consensus to allow contiguation"

/citation=[2]

1526..1855

/gene="SPBC27B12.02"

/note="synonym: SPBC30B4.10"

1526..1855

/gene="SPBC27B12.02"

/note="very hypothetical protein"

/codon\_start=1

/protein\_id="CA16897.1"

/db\_xref="GI:2853108"

/translation="MPLEKARAIEIAPNVFNTKIPNIOQFDALIKLERRRIPT ENQKRVYETELVTRFERGVKXDNHNPINLHLSRSKSLIAQSKLDLHNRVFLR NVRAVVM"

/complement(2261..3250)

/gene="SPBC27B12.03c"

/complement(2261..3250)

/gene="SPBC27B12.03c"

/note="C-5 sterol desaturase (predicted); involved in ergosterol biosynthesis (predicted); similar to S. cerevisiae ERG3"

/codon\_start=1

/protein\_id="CA16898.1"

/db\_xref="GI:2853109"

/db\_xref="GOA:O13666"

/db\_xref="UniProt/Swiss-Prot:O13666"

/translation="MDVVLQYADKYVDFPYGKIKESPDSSSPANTAVNSPTTLGLAE KNPFAITSGLDLRNVMROFTSLFITWIMGLSYFLSASFAYVYVFDREARRHPKF



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2005, 13:36:05 ; Search time 1739 Seconds  
(without alignments)  
11999.476 Million cell updates/sec

Title: US-10-018-105-3

Perfect score: 3525  
1 tatatataataataata.....tcaattattcaaatatgc 3525

Sequence: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Scoring table: Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues 8780412

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

1: Geneseq1980s:\*\n2: Geneseq1990s:\*\n3: Geneseq2000s:\*\n4: Geneseq2001as:\*\n5: Geneseq2001bs:\*\n6: Geneseq2002as:\*\n7: Geneseq2002bs:\*\n8: Geneseq2003as:\*\n9: Geneseq2003bs:\*\n10: Geneseq2003cs:\*\n11: Geneseq2003ds:\*\n12: Geneseq2004as:\*\n13: Geneseq2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3523	99.9	3525	3 AAC87954	AAC87954 Candida a
2	2769	78.6	2769	6 ABZ32388	ABZ32388 Candida a
3	263.8	7.5	684	6 ABW21722	ABW21722 Human ORF
4	179.6	5.1	1992	8 ABT19393	ABT19393 Aspergill
5	179.6	5.1	1992	8 ABT19393	ABT19393 Aspergill
6	169.4	4.8	860	8 ABZ51492	ABZ51492 Aspergill
7	168.2	4.8	537	2 ADR01374	ADR01374 A. gossyp
8	118	3.3	462	6 ABW76186	ABW76186 Human gly
9	117.6	3.3	8056	8 ABZ10246	ABZ10246 Haematopo
10	116.4	3.3	688	3 AAF14580	AAF14580 Aspergill
11	113.6	3.2	2048	8 ABT20615	ABT20615 Aspergill
12	113.6	3.2	2048	8 ABT18799	ABT18799 Aspergill
13	113.6	3.2	4048	8 ABT18205	ABT18205 Aspergill
14	113.6	3.2	4048	8 ABT20019	ABT20019 Aspergill
15	113.6	3.2	7990	13 ADR84473	ADR84473 Human ORF
16	107.8	3.1	321	6 ABZ22505	ABZ22505 Human ORF
17	106.8	3.0	4985	6 ABG75107	ABG75107 Anopheles
18	106.8	3.0	4985	10 ACF79720	ACF79720 Mosquito
19	106.2	3.0	575	3 AAF07924	AAF07924 Fusarium
20	94.4	2.7	969	6 ABQ39490	ABQ39490 Oligonuc1

21	94.4	2.7	969	6 ABQ39491	ABQ39491 Oligonuc1
22	92.6	2.6	8056	8 ABZ10246	ABZ10246 Haematopo
23	90	2.6	3996	6 AAD47004	AA47004 Plasmodiu
24	86.6	2.5	8056	8 ABZ10100	ABZ10100 Haematopo
25	85.8	2.4	2000	10 ADK62467	ADK62467 Diasease t
26	84.8	2.4	8056	8 ABZ10100	ABZ10100 Haematopo
27	84	2.4	7676	6 ABL34598	ABL34598 Human met
28	84	2.4	7676	6 ABL70409	ABL70409 Human met
29	84	2.4	7676	7 ADS99859	ADS99859 Bistulphit
30	81.2	2.3	4985	6 ABQ75107	ABQ75107 Anopheles
31	81.2	2.3	4985	10 ACF79720	ACF79720 Mosquito
32	79.8	2.3	6033	3 AAA70152	AAA70152 Plasmodiu
33	79.8	2.3	5859	13 ADS89440	ADS89440 Oligonuc1
34	79.8	2.3	9859	13 ADS89714	ADS89714 Oligonuc1
35	79.6	2.3	3931	9 ACC84661	ACC84661 P. falcip
36	78.2	2.2	8759	8 ABZ10238	ABZ10238 Haematopo
37	78.2	2.2	8759	8 ABZ09964	ABZ09964 Haematopo
38	78.2	2.2	8759	8 ABZ10092	ABZ10092 Haematopo
39	78.2	2.2	8759	8 ABZ10110	ABZ10110 Haematopo
40	78.2	2.2	8759	10 ADB54086	ADB54086 Pretreate
41	78.2	2.2	8759	10 ADB54214	ADB54214 Pretreate
42	78.2	2.2	8759	10 ADB84072	ADB84072 Human lym
43	78.2	2.2	8759	10 ADB84148	ADB84148 Human lym
44	78.2	2.2	8759	13 ADS89514	ADS89514 Oligonuc1
45	78.2	2.2	8759	13 ADS89240	ADS89240 Oligonuc1

## ALIGNMENTS

RESULT 1  
AAC87954  
ID AAC87954 standard; DNA; 3525 BP.

XX AC AAC87954;  
XX DT 06-MAR-2001 (first entry)  
XX DE Candida albicans CaLR1 nucleotide sequence.  
XX KM Candida albicans, yeast pathogen; identification; fungal; antifungal;  
XX KW CaLR5; CaLR1; CaLR24; diagnosis; fungicide; fungal infection; de.  
XX OS Candida albicans.  
XX PN WO200068420-A2.  
XX PD 16-NOV-2000.  
XX PF 05-MAY-2000; 2000WO-CA000533.  
XX PR 05-MAY-1999; 99US-0132878P.  
XX XX (MYCO-) MYCOTA BIOSCIENCES INC.  
XX PI Roemer T, Bussey H, Davison J;  
XX DR WPI; 2000-687652/67.  
XX DR P-PSDB; AAB36515.  
XX PT New DNA encoding essential proteins of Candida albicans, useful for  
XX PT diagnosing fungal infections and to screen for clinical or agricultural  
XX PT antifungal agents.  
XX PS Claim 3; Fig 2A; 76pp; English.  
XX PS The present sequence represents the fungus-specific Candida albicans gene  
XX CC CaLR1. The present invention describes the fungus-specific genes CaLR5,  
XX CC CaLR1 and CaLR24 isolated in the yeast pathogen C. albicans. The genes  
XX CC have antifungal and fungicide activity. The genes in C. albicans and are  
XX CC useful as drug targets. Fragments of them are useful as probes and  
XX CC primers for diagnosis of fungal infections, also as antisense and  
XX CC ribozyme agents. Proteins encoded by the genes are used to screen for







expressible selectable marker and modifying other allele by recombination, of a promoter replacement fragment with a heterologous promoter, so that expression of the second allele is regulated by the promoter. (M) is useful for constructing a strain of diploid fungal cells in which both alleles of a gene are modified. The diploid fungal cells having both alleles modified are useful for identifying a gene that is essential to the survival or growth of a fungus, a gene that contributes to the virulence and/or pathogenicity of a fungus, a gene that contributes to the resistance of a diploid fungus to an antifungal agent, an antifungal agent that inhibits the growth of a diploid fungus and for identifying a therapeutic agent for treatment of a mammalian disease. (M1) is useful for identifying a compound which modulates the activity of a gene product, preferably enzymatic activity, carbon compound catabolism, biosynthetic, transporter, transcriptional, translational, signal transduction, DNA replication and cell division activity. The method is useful for identifying a compound having the ability to inhibit growth or proliferation of C. albicans cells and for treating infection by C. albicans. The present sequence is that of an essential Candida albicans gene used in the method of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office

Sequence 2769 BP, 987 A; 434 C; 448 G; 900 T; 0 U; 0 Other;

Query Match 78.6%; Score 2769; DB 6; Length 2769;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2769; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

338 ATGTCGATGTAAGAAAGTTATTCATAATTCACCTACTATCAACCTTCTGATCT 397  
 1 ATGTCGATGTAAGAAAGTTATTCATAATTCACCTACTATCAACCTTCTGATCT 60  
 339 GATGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 457  
 61 GATGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
 458 GATGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 517  
 121 GATGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
 518 CAACAACATCATCAGAGATTCATCAGATGATGATGATGATGATGATGATGAT 577  
 181 CAACAACATCATCAGAGATTCATCAGATGATGATGATGATGATGATGATGAT 240  
 578 TCTTCATTAAGAAAGTCTATCTTACCGATTAAGATGATGATGATGATGATGAT 637  
 241 TCTTCATTAAGAAAGTCTATCTTACCGATTAAGATGATGATGATGATGATGAT 300  
 638 TATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 697  
 301 TATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
 698 TCATTAAGTAAAGATTTTATTTAAAGTAAATCTGACGACATTTCTATTAATCT 757  
 361 TCATTAAGTAAAGATTTTATTTAAAGTAAATCTGACGACATTTCTATTAATCT 420  
 758 ACTCAATCTGCAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGA 817  
 421 ACTCAATCTGCAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGA 480  
 818 AATCAAGAAAGAAATCTCAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGA 877  
 481 AATCAAGAAAGAAATCTCAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGA 540  
 878 GGTGCTAATTAATTTCTAATTAATTTGAAATGATTTAGTCCATGCAAAAATGAAA 937  
 541 GGTGCTAATTAATTTCTAATTAATTTGAAATGATTTAGTCCATGCAAAAATGAAA 600  
 938 ACTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 997  
 601 ACTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660

998 ATGTCGATGTAAGAAAGTTATTCATAATTCACCTACTATCAACCTTCTGATCT 1057  
 661 ATGTCGATGTAAGAAAGTTATTCATAATTCACCTACTATCAACCTTCTGATCT 720  
 1058 GCTGCTGTAAGAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1117  
 721 GCTGCTGTAAGAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGAT 780  
 1118 AGATCATCTCAAGAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1177  
 781 AGATCATCTCAAGAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGAT 840  
 1178 GTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1237  
 841 GTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
 1238 TATTTTCAAAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1297  
 901 TATTTTCAAAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
 1238 TCCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1357  
 961 TCCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1020  
 1358 AATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1417  
 1021 AATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080  
 1418 CCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1477  
 1081 CCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1140  
 1478 TCTTCAGAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1537  
 1141 TCTTCAGAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
 1538 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1597  
 1201 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260  
 1598 TTTGCTGAGCAAGAAATTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1657  
 1261 TTTGCTGAGCAAGAAATTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 1658 CTTTCCATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1717  
 1321 CTTTCCATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
 1718 CCAAGTCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1777  
 1381 CCAAGTCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
 1778 ACTGCTCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1837  
 1441 ACTGCTCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
 1838 ACTGCTCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1897  
 1501 ACTGCTCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
 1898 TATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1957  
 1561 TATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1620  
 1958 AATAAATGTTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2017  
 1621 AATAAATGTTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1680  
 2018 TCTCATCAGCAAAATGTTAAGAAAGTCTGATGATGATGATGATGATGATGATGATGAT 2077  
 1681 TCTCATCAGCAAAATGTTAAGAAAGTCTGATGATGATGATGATGATGATGATGATGAT 1740  
 2078 GCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2137

```

Db 1741 GGTGATGTTATGTTATGCTTATCGATGAAATTACGATGGTTGGCCCCCGTGAAT
Qy 2138 CATGGAATTTGAATGATGCTGATGCCCATTTGAGAGTCCGTTTCACTGTGAGAAATCT
Db 1801 CATGGAATTTGAATGATGCTGATGCCCATTTGAGAGTCCGTTTCACTGTGAGAAATCT
Qy 2198 GATTATTAGTATGTTTCAAAAGAAATGTTGATGAAATGAGAAAGAAAGTCACTGATTAAAG
Db 1861 GATTATTAGTATGTTTCAAAAGAAATGTTGATGAAATGAGAAAGAAAGTCACTGATTAAAG
Qy 2258 AGATTATTATGCTGATGAAAGTGTGATGCTTAAATGTTTGTCTAAAGATGTCAGAGAGA
Db 1921 AGATTATTATGCTGATGAAAGTGTGATGCTTAAATGTTTGTCTAAAGATGTCAGAGAGA
Qy 2318 GCTAATCTCTCTGCTGTTTATATCAAGTCAATTAATTAATTAATTAATTAATTAATTAAT
Db 1981 GCTAATCTCTCTGCTGTTTATATCAAGTCAATTAATTAATTAATTAATTAATTAATTAAT
Qy 2378 GCCCCACACACACACCTTATCTTATTAATTAATTAATTAATTAATTAATTAATTAATCT
Db 2041 GCCCCACACACACACCTTATCTTATTAATTAATTAATTAATTAATTAATTAATTAATCTT
Qy 2438 AATAGTTTGAAGAACTTCAACTGCTGAGAGAGTGAAGAGTGAAGAAATTAATTTGGTCCC
Db 2101 AATAGTTTGAAGAACTTCAACTGCTGAGAGAGTGAAGAGTGAAGAAATTAATTTGGTCCC
Qy 2498 AATCCAACTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
Db 2161 AATCCAACTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
Qy 2558 CAACAAACAACAACAACATGATGATCACTAACTTTTCCCATCCCGATCCGATCCAGTCCGA
Db 2221 CAACAAACAACAACAACATGATGATCACTAACTTTTCCCATCCCGATCCGATCCAGTCCGA
Qy 2618 AGAGCTGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
Db 2281 AGAGCTGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
Qy 2678 CAAGTTGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
Db 2401 CAAGTTGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
Qy 2798 AATGGGACAATGTTGTTCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
Db 2461 AATGGGACAATGTTGTTCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
Qy 2858 GTCCCTGATGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
Db 2521 GTCCCTGATGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
Qy 2918 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
Db 2581 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
Qy 2978 AATGGAAGACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
Db 2641 AATGGAAGACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
Qy 3038 AGAAGTTTGGTTTAAAAAATGATGATGATTAATTAATTAATTAATTAATTAATTAATTA
Db 2701 AGAAGTTTGGTTTAAAAAATGATGATGATTAATTAATTAATTAATTAATTAATTAATTA
Qy 3098 TATGAATTA 3106
Db 2761 TATGAATTA 2769

```

RESULT 3

```

ABN21722
ID ABN21722 standard; cDNA; 684 BP.
XX
AC ABN21722;
XX
DT 24-JUN-2002 (first entry)
XX
DE Human ORFX polynucleotide sequence SEQ ID NO:11921.
XX
KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; hemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis; gene; ss.
XX
OS Homo sapiens.
XX
PN MO200192523-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US010836.
XX
PR 30-MAY-2000; 2000US-0206132P.
XX
PR 29-AUG-2000; 2000US-0228716P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinketsu RA, Leach MD;
XX
DR WPI, 2002-106308/14.
XX
DR P-PSDB; ABP05970.
XX
PT Novel human polypeptides and polynucleotides useful for diagnosing,
XX preventing and treating cardiovascular disease, neurodegenerative,
XX hyperproliferative disorders and autoimmune disorders.
XX
PS Disclosure; SEQ ID NO 11921; 1037pp; English.
XX
SS
CC The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP0010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, hemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage. N.B. The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPD at ftp.wipo.int/pub/published_pat_sequences
XX
SQ Sequence 684 BP; 223 A; 105 C; 140 G; 216 T; 0 U; 0 Other;

```

Query Match 7.5%; Score 263.8; DB 6; Length 684;  
 Best Local Similarity 70.0%; Pred. No. 1.2e-36;  
 Matches 355; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

QY 1865 ACTGCTGAAAAAGTGAATTATTAAGTATTAATTTGTTGTTCCATCTTTGAA 1924  
 DB 1 ACGCGTGAAGAGGTAGAGCTTTTCAATCCATCTTGTCTGTTCCATCATTTGAA 60  
 QY 1925 GCTGATTAAGATCTGAAGATTTATTAAGACCGATTAATGTTTATTTGTTCCAT 1984  
 DB 61 AATGATTAAGATCTGAAGATTTATTAAGACCGATTAATGTTTATTTGTTCCAT 120  
 QY 1985 GATGATTAATTAAGTCAATTTTCCATTTTCCATTTCCATTTCCATTTCCATTT 2044  
 DB 121 TCAGGTCTTGAATCTTTTCAATTTTCCATTTTCCATTTTCCATTTTCCATTT 180  
 QY 2045 GTTCGCTAATTAAGATTAATTTGCTGATTTGCTGATTTGCTGATTTGCTGATTT 2104  
 DB 181 GTAAGACATTTACGATCTATGCTGATTTGCTGATTTGCTGATTTGCTGATTTGCT 240  
 QY 2105 GATGAATTTACCGATGTTTGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2164  
 DB 241 GATGATTTACCGATGTTTGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
 QY 2165 ATTAAGATTTACCGATGTTTGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2224  
 DB 301 ATTAAGATTTACCGATGTTTGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
 QY 2225 GGTGAATCAAGAAAGTCAATGCTTTAATGATTTAATGATTTAATGATTTAATG 2284  
 DB 361 GGTGAATCAAGAAAGTCAATGCTTTAATGATTTAATGATTTAATGATTTAATGAT 420  
 QY 2285 ATTAAGATTTACCGATGTTTGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2344  
 DB 421 ATCAAAATGTTTGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 QY 2345 CGTCAATTAATCTTCAACCAACCAACCA 2371  
 DB 481 TCTCAACTAATCACTGATTAATCA 507

RESULT 4  
 ABT19393  
 ID ABT19393 standard, DNA, 1992 BP.

AC ABT19393;  
 DT 16-APR-2003 (first entry)  
 DE Aspergillus fumigatus essential gene #1751.  
 KM Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;  
 KW cancer; contamination; biofilm; antibody; immune response; ds.  
 OS Aspergillus fumigatus.  
 PN WO200286090-A2.  
 PD 31-OCT-2002.  
 PF 23-APR-2002; 2002WO-US013142.  
 PR 23-APR-2001; 2001US-028697P.  
 PR 27-APR-2001; 2001US-028706P.  
 PR 05-JUN-2001; 2001US-0295890P.  
 PR 09-JUL-2001; 2001US-0303899P.  
 PR 31-AUG-2001; 2001US-0316362P.

PA (ELIT-) ELITRA PHARM INC.

PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;

XX MPI, 2003-093124/08.

DR New purified or isolated nucleic acids of essential genes of Aspergillus  
 PT fumigatus, useful for treating or preventing infections by A. fumigatus,  
 or for treating a non-infectious disease in a subject e.g. cancer.

XX Disclosure, Page: 175pp, English.

PS The invention relates to novel purified or isolated nucleic acids of  
 CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of  
 CC the invention are used to treat or prevent infections by a pathogenic  
 CC organism such as A. fumigatus, to treat a non-infectious disease in a  
 CC subject (e.g. cancer), to prevent or contain contamination of an object  
 CC by A. fumigatus, or to prevent or inhibit formation on a surface of a  
 CC biofilm comprising A. fumigatus. The polynucleotides are useful for  
 CC expressing recombinant protein for characterization, screening or  
 CC therapeutic use, as markers for host tissues in which the pathogenic  
 CC organisms invade or reside, for comparing with the DNA sequence of A.  
 CC fumigatus to identify duplicated genes or paralogues having the same or  
 CC similar biochemical activity and/or function, for comparing with DNA  
 CC sequences of other related or distant pathogenic organisms to identify  
 CC potential orthologous essential or virulence genes, for selecting and  
 CC making oligomers for attachment to a nucleic acid array for examination  
 CC of expression patterns, for raising anti-protein antibodies, as an  
 CC antigen to raise anti-DNA antibodies or to elicit another immune  
 CC response, and for identifying polynucleotides encoding the other protein  
 CC with which binding occurs or to identify inhibitors of the binding  
 CC interaction. The polypeptides may be used to raise antibodies or to  
 CC elicit immune response, as a reagent in assays designed to quantitatively  
 CC determine levels of the protein in biological fluids, as a marker for  
 CC host tissues in which pathogenic organism invade or reside, and to  
 CC isolate correlative receptors or ligands in the case of virulence  
 CC factors. This polynucleotide sequence represents one of the essential  
 CC genes of Aspergillus fumigatus of the invention

SQ Sequence 1992 BP; 458 A; 509 C; 525 G; 500 T; 0 U; 0 Other;

Query Match 5.1%; Score 179.6; DB 8; Length 1992;

Best Local Similarity 56.6%; Pred. No. 7e-22;

Matches 376; Conservative 0; Mismatches 279; Indels 9; Gaps 2;

QY 1659 TTTTCATCTGATGATCAGAGAAACTATTCATGCCCGATATTCATATTAATCAAC 1718  
 DB 959 TTTTCTCATCCGAGTCCGAGACACTGTCATGCGCGGAATCGGAGCACTTGCTGCC 1018  
 QY 1719 CAGGTCATCTGTTGAGATTTATTTAGAAATG-----TGAAGAACTTGCTGCTG 1772  
 DB 1019 CTGGGATATCTTCCGGATCTCTTTCAACTTGGCCGAGGAGGTGTGGTGGTGG 1078  
 QY 1773 ATTTGACTTGTCTACTGATTCGGAATGAAATGTTGGCCAAAGCTTTGGTATTCATC 1832  
 DB 1079 ACGTACTAATTCACAGAGAGAGAGAGTGTCTGCACTCTCCGGGCAATTTTCATCATT 1138  
 QY 1833 CTTTAACCTGTAAGATATTCGAATGCAAGAACTGTGAAGAAAGTTGAATTTAATAA 1892  
 DB 1139 CACTGACTACGGAAGATATCTTGACTCAAGAGCCCGTGAAGAGGTGAGCTTTCAAGC 1198  
 QY 1893 GTTATTTATTTGTTGTTTCCATCTT---TGAAGCTGATTAAGAAATCTGAAGTTATT 1949  
 DB 1199 AATATTAATCTTGTCTGCTTCCGACATTCATGCTGACAGACAGACAGACAGACGCTTCA 1258  
 QY 1950 TAGAACGATTAATGTTATATATGTTGTTTCCATGATGATTAATTAAGTTCCATTTT 2009  
 DB 1259 TGAAGCCGCTCACTTCTACATGTTGTTTCCGAGAGGCTGCTCTCATTTTCATTTCA 1318  
 QY 2010 CACCAATTTCTATCCAGCAATATTTAGAGAGAGTGTCTCAATTTAGAGATTAATGCTG 2069  
 DB 1319 CCGAAGATCCACAGCGGCAATATGTCAGAGAGAAATTTGGGAAGCTTGATTAATGTTG 1378  
 QY 2070 ATGTTAGCTGATTTGTTATGTTATGCTTAATCGATGAATTAACGATGTTTGGCC 2129  
 DB 1379 CCGTCAGTAGAGCTGATTCGTTATGCTTAATGATTAAGACATTTGTTATGTTTGGCC 1438  
 QY 2130 CCGTATTCATGATTAATGAATATGATGAAGTGCATGATGAAGATGCGCTTTTCACTGCTA 2189  
 DB 1439 CTGTTCATCCGAGATTTAGATTCGATTTGAGGCAATTTGAGGACCTGTTTTCATCGGCG 1498  
 QY 2190 GAGATTAATGATTTAGTATGTTTACAAAGAAATTTGTAATCAAGAGAAAGTCAATGA 2249

Db 1499 GGGTCGACGATTTTGAATCTCTTACCGCGAATCGGTGGCCCTTGGTAAGAAGTCATGA 1558  
 QY 2250 CTTTAATGAGATTATATACGTAAGCGATGATCTTAAATGTTTGTATAAAGATGTC 2309  
 Db 1559 GCTTATATGGCCCTTCTAGGTGGCAAGAGATGTCAATTCGCGTTTCTTCAAGCGTTGCA 1618  
 QY 2310 AAGA 2313  
 Db 1619 ACGA 1622

RESULT 5  
 AB21213  
 ID AB21213 standard; DNA; 1992 BP.  
 AC AB21213;  
 XX  
 DT 16-APR-2003 (first entry)  
 DE Aspergillus fumigatus essential gene #3571.  
 XX  
 KM Fungicide; cytosstatic; essential gene; Aspergillus fumigatus; infection;  
 KM cancer; contamination; biofilm; antibody; immune response; ds.  
 XX  
 OS Aspergillus fumigatus.  
 XX  
 PN WO200286090-A2.  
 PD 31-OCT-2002.  
 PF 23-APR-2002; 2002WO-US013142.  
 PR 23-APR-2001; 2001US-0285697P.  
 PR 27-APR-2001; 2001US-0287066P.  
 PR 05-JUN-2001; 2001US-0295890P.  
 PR 09-JUL-2001; 2001US-0303899P.  
 PR 31-AUG-2001; 2001US-0316362P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;  
 XX  
 DR WPI; 2003-093124/08.  
 XX  
 PT New purified or isolated nucleic acids of essential genes of Aspergillus  
 PT fumigatus, useful for treating or preventing infections by A. fumigatus,  
 PT or for treating a non-infectious disease in a subject e.g. cancer.  
 XX  
 PS Disclosure; Page; 175pp; English.  
 XX  
 CC The invention relates to novel purified or isolated nucleic acids of  
 CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of  
 CC the invention are used to treat or prevent infections by a pathogenic  
 CC organism such as A. fumigatus, to treat a non-infectious disease in a  
 CC subject (e.g. cancer), to prevent or contain contamination of an object  
 CC by A. fumigatus, or to prevent or inhibit formation on a surface of a  
 CC biofilm comprising A. fumigatus. The polynucleotides are useful for  
 CC expressing recombinant protein for characterisation, screening or  
 CC therapeutic use, as markers for host tissues in which the pathogenic  
 CC organisms invade or reside, for comparing with the DNA sequence of A.  
 CC fumigatus to identify duplicated genes or paralogues having the same or  
 CC similar biochemical activity and/or function, for comparing with DNA  
 CC sequences of other related or distant pathogenic organisms to identify  
 CC potential orthologous essential or virulence genes, for selecting and  
 CC making oligomers for attachment to a nucleic acid array for examination  
 CC of expression patterns, for raising anti-protein antibodies, as an  
 CC antigen to raise anti-DNA antibodies or to elicit another immune  
 CC response, and for identifying polynucleotides encoding the other protein  
 CC with which binding occurs or to identify inhibitors of the binding  
 CC interaction. The polypeptides may be used to raise antibodies or to  
 CC elicit immune response, as a reagent in assays designed to quantitatively  
 CC determine levels of the protein in biological fluids, as a marker for

CC host tissues in which pathogenic organism invade or reside, and to  
 CC isolate correlative receptors or ligands in the case of virulence  
 CC factors. This polynucleotide sequence represents one of the essential  
 CC genes of Aspergillus fumigatus of the invention  
 XX  
 SQ Sequence 1992 BP; 458 A; 509 C; 525 G; 500 T; 0 U; 0 Other;  
 Query Match 5.1%; Score 179.6; DB 8; Length 1992;  
 Best Local Similarity 56.6%; Pred. No. 76-22;  
 Matches 376; Conservative 0; Mismatches 279; Indels 9; Gaps 2;

QY 1659 TTTTCATTCTGGAATCAGAAAGAACTATTCATGCCCCGATATTCATCTAGTATCAC 1718  
 Db 959 TTTTCTCATCCGAGTCCAGACACTGTGATGCGGCGAATCGGGGCACTTGTGCTCC 1018  
 QY 1719 CAGTCATCTGTTGGAATTTATTTAAGAAATG-----TGAAGAACTTGTTGTTG 1772  
 Db 1019 CTGGTGAATACCTTTCGGGATCTCTTCAACTTGGCCCGAGGAGGTGTGTGTGTGG 1078  
 QY 1773 ATTGACTTGTCTCACTGATTCGAAATGAATAATGTTGGCCAAACATTTGTTGATTC 1832  
 Db 1079 ACGTACTCAATCCAAAG 1138  
 QY 1833 CTTTAACCTGCTGAAGATATTCGAATGCAAGAACTGTGAAAAGTTGAATTTTAA 1892  
 Db 1139 CACTGACTACGGAAGATATCTTGACTCAAGAAAGCCGTGAAAAGGTGAGCTTTTCAAGC 1198  
 QY 1893 GTTATTTATTTTGTGTTTCCACTT---TGAAGCTGATTAAGATTCGAAGATTA 1949  
 Db 1199 AATATTACTTGTCTGCTTCCGACATCTTACACTCGACACAGACAGACAGCTTCA 1258  
 QY 1950 TAGAACCGATTAATGTTATATTTGTTTTCATGATGATATTAAGTTCCATTTT 2009  
 Db 1259 TGGAGCCCGTCACTTCTACATGTTGTTTCCGAGCGGTCTCTCATTTCTCATTA 1318  
 QY 2010 CACCAATTTCTCATCAGCAAAATGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2069  
 Db 1319 CCGAAGATCCACACCGGCAATGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1378  
 QY 2070 ATGTTAGTCTGATGTTGTTATGTTATGCTTATGATGAATTAACCATGTTTGGCC 2129  
 Db 1379 CCTCATAGTAGTACGAGATCTGTTATGCCATATGACACATTTGTATATGTTTGGGC 1438  
 QY 2130 CCGATTCATGGAATTAATTAATGAAGCTGATGCAATGGAATGCGTTTCACTGCTA 2189  
 Db 1439 CTGTATCCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1498  
 QY 2190 GAGATACGATTTTGTAGTATGTTACAAAGATTTGTAATCAAGAGAGAGAGAGAGAG 2249  
 Db 1499 GCGTCGACGATTTTGAATCTCTTACCGCGAATCGGTGGCCTTGTAAAGAGTCATGA 1558  
 QY 2250 CTTTAATGAGATTATATACGTAAGCGATGATCTTAAATGTTTGTATAAAGATGTC 2309  
 Db 1559 GCTTATATGGCCCTTCTAGGTGGCAAGAGATGTCAATTCGCGTTTCTTCAAGCGTTGCA 1618  
 QY 2310 AAGA 2313  
 Db 1619 ACGA 1622

RESULT 6  
 AB251492  
 ID AB251492 standard; cDNA; 860 BP.  
 XX  
 AC AB251492;  
 XX  
 DT 28-MAR-2003 (first entry)  
 DE Aspergillus oryzae polynucleotide SEQ ID NO 605.  
 XX  
 KM Aspergillus oryzae; fermentation; fungus; industrial; EST;  
 KM expressed sequence tag; gene; ss.  
 XX





PI Lewin A, Lipocher E, Maier S, Model F, Mueller V, Otto T, Pelet C;  
PI Schwope I, Ziebarth H;  
XX MPI, 2003-018942/01.

PT Detecting and differentiating between hematopoietic cell proliferative  
PT disorders, comprises contacting a target nucleic acid with a reagent that  
PT distinguishes between methylated and non-methylated CpG dinucleotides.

PS Claim 28; SEQ ID NO 386; 117bp; English.

XX The present invention describes a method for detecting and  
CC differentiating between hematopoietic cell proliferative disorders  
CC associated with at least 1 gene and/or their regulatory regions in a  
CC subject. The method comprises contacting a target nucleic acid in a  
CC biological sample obtained from the subject with at least 1 reagent,  
CC which distinguishes between methylated and non-methylated CpG  
CC dinucleotides within the target nucleic acid. AB209861 to AB211118  
CC represent specifically claimed nucleotide sequences from the present  
CC invention. Oligonucleotides from the present invention can be used: for  
CC differentiating between healthy hematopoietic cells and proliferative  
CC disorder hematopoietic cells; for differentiating between acute  
CC lymphocytic leukemia and acute myelogenous leukemia; as probes for  
CC determining the cytosine methylation state and/or single nucleotide  
CC polymorphisms (SNPs) of hematopoietic cell proliferation disorder  
CC related sequences and their complements; and as primers for the  
CC amplification of hematopoietic cell proliferation disorder related DNA  
CC sequences. The nucleotide sequences from the present invention can also  
CC be used for detecting a predisposition to, differentiation between  
CC subtypes, diagnosis, prognosis, treatment and/or monitoring of  
CC hematopoietic cell proliferative disorders. The present method enables a  
CC highly specific classification of hematopoietic cell proliferative  
CC disorders allowing for improved and informed treatment of patients

SQ Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;

Query Match 3.3%; Score 117.6; DB 8; Length 8056;  
Best Local Similarity 42.2%; Pred. No. 5.8e-11;  
Matches 1426; Conservative 0; Mismatches 1911; Indels 43; Gaps 12;

QY 169 TTAATATTAAGATNTCCATTTTTTTTTTCCAGCTAGTAAATTTATTTGGTGT 228  
DB TTTTATTTATTTTATTTTATTTTATTTTAAAAAATTTAAATTTTATTTTAAAT 1463  
QY 229 CTAACTATATATATATTTTACAGAAATTCCTCAATATTAATTAATTAATTA 288  
DB 1464 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1522  
QY 289 ATATATTAAGATATATCTCCCTTTGTTTTTTTTTTCTCCAGCCATGTCGATNG 348  
DB 1523 ATGTAAAAAATTTTAAATTAATTTTAAATTAATTAATTAATTTAAAT 1582  
QY 349 TGAATGTATTTTCAAAATTCATCTATGACCTATCTCTAGTCGATGAAGTAT 408  
DB 1583 TTTAATTTATTTTATTTTATTTTAAATTTTAAATTAATTAATTAATTTAAAT 1642  
QY 409 GGATGATCATAGAAATCAATCATATGATTTGTCATAGTATGATGAAGATGAT 468  
DB 1643 AAAAAAATTTATTTATTAATTAATTAATTAATTTATTTATTAATTAATTAATTA 1702  
QY 469 GGAATTAATAATCAAGATTGAATCGAAGTTGTAAGCGAAAAACAACAACATCA 528  
DB 1703 -ATTATTAATTAATTAATTAATTAATTTAAATTTAAATTAATTAATTAATTA 1759  
QY 529 TCAAGAGATTACATGATGATGCTAAACCATGCTGTAATCTGGTCTTCAATTA 588  
DB 1760 -AATTTTAAATTAATTAATTAATTAATTAATTTATTTATTTATTAATTAAT 1817  
QY 589 GAAAAATATCTATCTTACGATTAAGATGAATTAACAACCTATGATGTTCTGCTG 648  
DB 1818 AAAATTAATTTTGTTTTAAAAAATTTAAAAAATTTAAATTTATTAATTAATTA 1877  
QY 649 TGATGATCTATTAACAAGGTCACAAAAATCGTAATTAACAATGATTCATTAAGTAA 708

DB 1878 TTTTATTTTATTTTATTAATTAATTAATTAATTTATTAATTAATTAATTAATTAATTA 1937  
QY 709 AGATTTTATTTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 768  
DB 1938 ATATTAATTAATTAATTAATTAATTTTAAATTAATTAATTAATTAATTTATTTATTA 1997  
QY 769 A--ATTCATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGA 825  
DB 1998 ATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTATTTT 2057  
QY 826 AAGAAATTCGAATTTGAAAATTTACCTCATTAATTAATTAATTAATTAATTAATTAATTA 885  
DB 2058 AAAAAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2116  
QY 886 TAAATTCATTAATTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 945  
DB 2117 TAAATTTATTAATTTTATTTATTTATTAATTAATTAATTAATTAATTAATTAATTAAT 2176  
QY 946 TAGTAAGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1005  
DB 2177 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTAA 2236  
QY 1006 TACAACCTTGCTGCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1065  
DB 2237 TAAAAATTTATTTATTAATTAATTAATTAATTAATTAATTAATTTTATTTATTTATTT 2296  
QY 1066 TA---GAGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1121  
DB 2297 TATTTTAAATTAATTAATTAATTAATTTTGTATTTAAATTTATTTATTTATTTATTTAA 2356  
QY 1122 CATCTCAAGAACTGAAGAGATGTTGTTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 1181  
DB 2357 TAAATTTTATTTTATTAATTAATTAATTAATTAATTTATTTATTTTATTTTATTTTAA 2416  
QY 1182 ATGAATTTGATTTGATGAATTTGATGAATTTATTAAGAAAGAAAGAGAGAGAGATTTAT 1241  
DB 2417 AAATTAATTTATTTATTTTAAATTAATTTATTTTATTTTATTTATTTATTTATTTATTT 2476  
QY 1242 TACAATAAATCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1301  
DB 2477 TATTTTGTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2536  
QY 1302 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1361  
DB 2537 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2596  
QY 1362 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1421  
DB 2597 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2656  
QY 1422 AAAATTAATTTTAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1481  
DB 2657 TAAATTTTAAATTAATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 2716  
QY 1482 CAGAGAAATTTATGAATTAAGACTAAACAACAACCTTACCAATATATGATCAAT 1541  
DB 2717 TAAAAATTAATTTGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTTGTTTAT 2776  
QY 1542 TATCATTAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCT 1601  
DB 2777 TTTAATTAATTAATTAATTAATTTTGTTAATTAATTAATTTGTTAATTAATTAATTAATTT 2836  
QY 1602 GTGAGCAAGAAAT--TTCTGATGGAATTAATGAGGTTCAATTAACCTGATGATATTTTCACTT 1660  
DB 2837 TTTATTAATTAATTAATTAATTTTGAATTTATTTATGAAGAAAAATTTTGTGAAAAAATA 2896  
QY 1661 TTCAATTCGATCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1720  
DB 2897 AAAATTTGAAATGAAGAAAAAATGTTAAATTTTATGAAGAAATGTTAAATTTTAAAT--GA 2953  
QY 1721 GGTCAATCTGTTGCAATTTATTTAAGAAATGCTGAAGAAATCTGCTGTTAGATTTGACT 1780









```

Db 959 TTTTCTCATCCGAGTCCAGACACTGTGATCGCGCGGAACTGTGCTCC 1018.
Qy 1719 CAGGTCAATCTGTGGAGATTATTTAGAAATG-----TGAAGAACTGGTGTTAG 1772
Db 1019 CTGGGATACCTTCCGGAGTCTTTCACCTTGGCCCGAGAGAGTGTGTGTGG 1078
Qy 1773 ATTGACTGTCTTACTGATTCGAAATGAAATGTTGGCCAAACATTTGGTATTCATC 1832
Db 1079 ACGTACTCAATCCACAGAGAGAGAGAGTTGCTGCACTCTCGGGGCAATTTTCATCCATC 1138
Qy 1833 CTTTACTGCTGAAGATATTCGAATGCAAGAACTGTGAAAAAGTTGAATTTTAA 1892
Db 1139 CACTGACTACCGGAAGATATCTTGAACCAAGCCCGTGAAGAGTTCAGCTTTCAAGC 1198
Qy 1893 GTTATTTATTTGTTGTTTCCACTT---TGAAGCTGATTAAGATCTGAAGTTAT 1949
Db 1199 AATATTACTTGTCTGCTTCGGAATCTTACACTGACACAGACAGAACCTTCA 1258
Qy 1950 TAGAACCGATAAATGTTATATTTGTTTCCATGATGTTATTAAGTTCCATTTT 2009
Db 1259 TGGAGCCCGTCACTTCAATGATGTTTTCGGAGCGGTCTCTCATTTCTCATCA 1318
Qy 2010 CACCAATTTCTCATCCAGCAATGTTAAGAGAGTTGTCATTTGAGATTAATGTCG 2069
Db 1319 CCGAATATCCACACGCGCAATGTCAGGAAGAGATTTGGAAAGCTTGATTAATGTCG 1378
Qy 2070 ATGTTAGTCTGATGTTGTTATGTTATGCTTAAATC----- 2104
Db 1379 CCTCAGTAGTACGATGATGTTATGTCATGATTAAGTCAAGAGCGTGATTTCTAA 1438
Qy 2105 -----GATGAAATTCAGATGTTGTTGCCCCCGT 2133
Db 1439 TGAGACCCCTAGTCTTCAACGTGGCAAGTACACATGTTGATAGTTTGGGCTGT 1498
Qy 2134 GATTCAGTAATGTAATGTAAGCTGATGCCATTTGAAGATGCGTTTCACTGCTAAGA 2193
Db 1499 CATCCGAGATTTGATGATGTAATGTAAGCAATTTGAAGCACTCGTTTCACTGCGCGT 1558
Qy 2194 TACTGATTTTGTAGTATGTTACAAAGAAATGTTGTAATCAGAAAGAAAGTATGACTT 2253
Db 1559 CGACGATTTTGAATCTTCTTACCGCGAATCGTGCGCTTCTTAAGAAAGTATGACTT 1618
Qy 2254 AATGATTTATATCAGGTAAGCTGATGTCATTAATAATGTTGTAAGAAAGTCAAGA 2313
Db 1619 GATGCGCTTTAGTGTGGCAAGCAAGATGTCATTTGCGGTTTCTCAAGCGTTCAACGA 1678

```

## RESULT 12

ABT18799 standard; DNA, 2048 BP.

```

ID ABT18799 standard; DNA, 2048 BP.
XX ABT18799;
XX AC
XX DT 16-APR-2003 (first entry)
XX DE Aspergillus fumigatus essential gene #1157.
XX KW Fungicide; cytosolic; essential gene; Aspergillus fumigatus; infection;
XX OS cancer; contamination; biofilm; antibody; immune response; ds.
XX OS Aspergillus fumigatus.
XX EN WO200286090-A2.
XX PD 31-OCT-2002.
XX PF 23-APR-2002; 2002MO-US013142.
XX PR 23-APR-2001; 2001US-0285697P.
XX PR 27-APR-2001; 2001US-0287066P.
XX PR 05-JUN-2001; 2001US-0295890P.
XX PR 09-JUL-2001; 2001US-0303899P.
XX PR 31-AUG-2001; 2001US-0316362P.

```

```

XX (ELIT-) ELITRA PHARM INC.
XX Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
XX WPI, 2003-093124/08.
XX New purified or isolated nucleic acids of essential genes of Aspergillus
XX fumigatus, useful for treating or preventing infections by A. fumigatus,
XX or for treating a non-infectious disease in a subject e.g. cancer.
XX Disclosure; Page; 175pp; English.
XX
XX The invention relates to novel purified or isolated nucleic acids of
XX essential genes of Aspergillus fumigatus. The isolated nucleic acids of
XX the invention are used to treat or prevent infections by a pathogenic
XX organism such as A. fumigatus, to treat a non-infectious disease in a
XX subject (e.g. cancer), to prevent or contain contamination of an object
XX by A. fumigatus, or to prevent or inhibit formation on a surface of a
XX biofilm comprising A. fumigatus. The polynucleotides are useful for
XX expressing recombinant protein for characterisation, screening or
XX therapeutic use, as markers for host tissues in which the pathogenic
XX organisms invade or reside, for comparing with the DNA sequence of A.
XX fumigatus to identify duplicated genes or paralogues having the same or
XX similar biochemical activity and/or function, for comparing with DNA
XX sequences of other related or distant pathogenic organisms to identify
XX potential orthologous essential or virulence genes, for selecting and
XX making oligomers for attachment to a nucleic acid array for examination
XX of expression patterns, for raising anti-protein antibodies, as an
XX antigen to raise anti-DNA antibodies or to elicit another immune
XX response, and for identifying polynucleotides encoding the other protein
XX with which binding occurs or to identify inhibitors of the binding
XX interaction. The polypeptides may be used to raise antibodies or to
XX elicit immune response, as a reagent in assays designed to quantitatively
XX determine levels of the protein in biological fluids, as a marker for
XX host tissues in which pathogenic organism invade or reside, and to
XX isolate correlative receptors or ligands in the case of virulence
XX factors. This polynucleotide sequence represents one of the essential
XX genes of Aspergillus fumigatus of the invention
XX
XX Sequence 2048 BP; 473 A; 520 C; 540 G; 515 T; 0 U; 0 Other;
XX
XX Query Match 3.2%; Score 113.6; DB 8; Length 2048;
XX Best Local Similarity 52.2%; Pred. No. 2.3e-10;
XX Matches 376; Conservative 0; Mismatches 279; Indels 65; Gaps 3;
XX
XX 1659 TTTTCCATTCGATCAGAGAGAACTATTCATGCCCGGATATTCATTAATGATCAG 1718
XX 959 TTTTCTCATCCGAGTCCAGACACTGTGATCGCGCGAAGTGGGGAACCTTGTCTCC 1018
XX 1719 CAGGTCAATCTGTGGAGATTATTTAGAAATG-----TGAAGAACTGGTGTTAG 1772
XX 1019 CTGGGATACCTTCCGGAGTCTTTCACCTTGGCCCGAGAGAGTGTGTGTGG 1078
XX 1773 ATTGACTGTCTTACTGATTCGAAATGAAATGTTGGCCAAACATTTGGTATTCATC 1832
XX 1079 ACGTACTCAATCCACAGAGAGAGAGAGTTGCTGCACTCTCGGGGCAATTTTCATCCATC 1138
XX 1833 CTTTACTGCTGAAGATATTCGAATGCAAGAACTGTGAAAAAGTTGAATTTTAA 1892
XX 1139 CACTGACTACCGGAAGATATCTTGAACCAAGCCCGTGAAGAGTTCAGCTTTCAAGC 1198
XX 1893 GTTATTTATTTGTTGTTTCCACTT---TGAAGCTGATTAAGATCTGAAGTTAT 1949
XX 1199 AATATTACTTGTCTGCTTCGGAATCTTACACTGACACAGACAGAACCTTCA 1258
XX 1950 TAGAACCGATAAATGTTATATTTGTTTCCATGATGTTATTAAGTTCCATTTT 2009
XX 1259 TGGAGCCCGTCACTTCAATGATGTTTTCGGAGCGGTCTCTCATTTCTCATCA 1318
XX 2010 CACCAATTTCTCATCCAGCAATGTTAAGAGAGTTGTCATTTGAGATTAATGTCG 2069
XX 1319 CCGAATATCCACACGCGCAATGTCAGGAAGAGATTTGGAAAGCTTGATTAATGTCG 1378

```



XX AB220019;  
 AC  
 XX 16-APR-2003 (first entry)  
 DT  
 XX Aspergillus fumigatus essential gene #2377.  
 DE  
 XX Fungicide; cytosolic; essential gene; Aspergillus fumigatus; infection;  
 KW cancer; contamination; biofilm; antibody; immune response; ds.  
 XX  
 OS Aspergillus fumigatus.  
 XX  
 EN WO200286090-A2.  
 XX  
 PD 31-OCT-2002.  
 XX  
 PE 23-APR-2002; 2002WO-US013142.  
 XX  
 PR 23-APR-2001; 2001US-0285697P.  
 PR 27-APR-2001; 2001US-0287066P.  
 PR 05-JUN-2001; 2001US-0295890P.  
 PR 09-JUL-2001; 2001US-0303899P.  
 PR 31-AUG-2001; 2001US-0316362P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 PI Jiang B, Tishkoff D, Zamudio C, Broshkin AM, Hu W, Lemieux SM;  
 XX WPI; 2003-093124/08.  
 DR  
 PT New purified or isolated nucleic acids of essential genes of Aspergillus  
 PT fumigatus, useful for treating or preventing infections by A. fumigatus,  
 PT or for treating a non-infectious disease in a subject e.g. cancer.  
 XX  
 XX Disclosure; Page; 175pp; English.  
 XX  
 CC The invention relates to novel purified or isolated nucleic acids of  
 CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of  
 CC the invention are used to treat or prevent infections by a pathogenic  
 CC organism such as A. fumigatus, to treat a non-infectious disease in a  
 CC subject (e.g. cancer), to prevent or contain contamination of an object  
 CC by A. fumigatus, or to prevent or inhibit formation on a surface of a  
 CC biofilm comprising A. fumigatus. The polynucleotides are useful for  
 CC expressing recombinant protein for characterization, screening or  
 CC therapeutic use, as markers for host tissues in which the pathogenic  
 CC organisms invade or reside, for comparing with the DNA sequence of A.  
 CC fumigatus to identify duplicated genes or paralogues having the same or  
 CC similar biochemical activity and/or function, for comparing with DNA  
 CC sequences of other related or distant pathogenic organisms to identify  
 CC potential orthologous essential or virulence genes, for selecting and  
 CC making oligomers for attachment to a nucleic acid array for examination  
 CC of expression patterns, for raising anti-protein antibodies, as an  
 CC antigen to raise anti-DNA antibodies or to elicit another immune  
 CC response, and for identifying polynucleotides encoding the other protein  
 CC with which binding occurs or to identify inhibitors of the binding  
 CC interaction. The polypeptides may be used to raise antibodies or to  
 CC elicit immune response, as a reagent in assays designed to quantitatively  
 CC determine levels of the protein in biological fluids, as a marker for  
 CC host tissues in which pathogenic organism invade or reside, and to  
 CC isolate correlative receptors or ligands in the case of virulence  
 CC factors. This polynucleotide sequence represents one of the essential  
 CC genes of Aspergillus fumigatus of the invention  
 CC  
 XX  
 SO Sequence 4048 BP; 925 A; 1043 C; 960 G; 1120 T; 0 U; 0 Other;

Query Match 3.2%; Score 113.6; DB 8; Length 4048;  
 Best Local Similarity 52.2%; Pred. No. 2.6e-10;  
 Matches 376; Conservative 0; Mismatches 279; Indels 65; Gaps 3;

QY 1659 TTTTCCATTTGGAATCAAGAAATATTCATGCCCCGAGTATTCATCATTAATGATAC.1718  
 DB 1959 TTTTCTCATCGAGTCCGAGACACTGTGCAATGCGCGGAACTGGGGAGACCTTGCTCC 2018

QY 1719 CAGTCAATCTGTGAGATTTATTTAGAAATGC-----TGAGAAACTGTGCTTAG 1772  
 DB 2019 CTGGTGAATATCTTCGCGGATCTTTCAACTGTGCGCCGAGAGGTGTGTGTGG 2078  
 QY 1773 ATTGACTGTTCCTACGATTCGAAATGAAATGTTGGCCAAAGCATTTGGTATTCATC 1832  
 DB 2079 ACGTACTCAATCCAAACAGAGAGAAAGTGTGCACTTCCGCGGATTTTCATCCATC 2138  
 QY 1833 CTTTAACTGCTGAAGATATTCGAATGCAAGAACTCGTGAAGAAAGTTGAATTAATTA 1892  
 DB 2139 CACTGACTACGGAAGATATCTTGACTCAAGAAAGCCGGAAGAGTCCAGCTTTTCAAGC 2198  
 QY 1893 GTTATTTATTTGTTGTTTCCATCTT---TGAAGTGAATTAAGATCTGAAGATTA 1949  
 DB 2199 AATATTAATCTTGTCTGCTTCGAGCAATCTATACGCTGACAGACAGACGACCTTCA 2258  
 QY 1950 TAGAACCGATTAAGTTTATATTTGTTTCCATGATGATATTAACGTTCCATTTT 2009  
 DB 2259 TGAAGCCCGTCAACTTCTACATGTTGTTTCCGAGCGGTCTCTCATTTCTCATTTCA 2318  
 QY 2010 CACCAATTTCTCATCCAGCAATGTTAGAGAGAGTTCATATTTAGAGATTAATGTCG 2069  
 DB 2319 CCGAATATCCACACGCGCAATGTCAGAGAGAAATTTGGAGCTTGTGATTAATG 2378  
 QY 2070 ATGTTAGTCTGATTTGTTATGTTATGCTTAATC----- 2104  
 DB 2379 CCTCAGTAGATCTGATCTGTTATATGCTATGATGATAGTCAAGAGCGGTGATTTCTAA 2438  
 QY 2105 -----GATGAATTAACGATGTTTCCCGCGT 2133  
 DB 2439 TGAGACCGTCTAGTCTTAACGTTGCGACAGTACGACATTTGTATGTTTGGCCCTGT 2498  
 QY 2134 GATTCATGGAATTTGAATTTGAAGCTGATGATGGAATGCGGTTTCACTGCTAGAGA 2193  
 DB 2499 CATCCGAGATTTGATGATCGAATCTGAGGCAATTTGAGGACCTCGTTTCATCGCCGCGT 2558  
 QY 2194 TACTGATTTTAGTAGTAGTTATGTTACAAAGATTTGTCATCAAGAAAGTCAATGACTTT 2253  
 DB 2559 CGAGATTTTGAATTCCTTTACCGCGAATCGGTGCGCTTGTAAGAGTATAGCTT 2618  
 QY 2254 AATGAGATTTATTCAGGTTAAAGCTGATGTCATTAATGTTTGTAAAGATGTCAGA 2313  
 DB 2619 GATGCGCTTCTAGGTGCAAGCAAGATGTCATTCGCGGTTTCTCAAGCGTTGCAAGCA 2678

RESULT 15  
 ADR84473  
 ID ADR84473 standard; DNA; 7990 BP.  
 XX  
 AC ADR84473;  
 XX  
 DT 04-NOV-2004 (first entry)  
 DE  
 XX Aspergillus fumigatus essential gene genomic sequence #284.  
 KW Fungicide; Aspergillus fumigatus infection; Farmer's lung disease;  
 KW drug screening; ds.  
 XX  
 OS Aspergillus fumigatus.  
 XX  
 EN WO2004067709-A2.  
 XX  
 PD 12-AUG-2004.  
 XX  
 PE 16-JAN-2004; 2004WO-US001099.  
 XX  
 PR 17-JAN-2003; 2003US-0441281P.  
 PR 13-JUN-2003; 2003US-0478196P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 PA (ELIT-) ELITRA PHARM LTD.  
 PI Jiang B, Hu W, Lemieux S, Roemer T;







CURRENT APPLICATION NUMBER: US/09/248,796A  
 CURRENT FILING DATE: 1999-02-12  
 PRIOR APPLICATION NUMBER: US 60/074,725  
 PRIOR FILING DATE: 1998-02-13  
 PRIOR APPLICATION NUMBER: US 60/096,409  
 PRIOR FILING DATE: 1998-08-13  
 NUMBER OF SEQ ID NOS: 28208  
 SEQ ID NO 6327  
 LENGTH: 507  
 TYPE: DNA  
 ORGANISM: Candida albicans  
 US-09-248-796A-6327

Query Match 12.0%; Score 423.8; DB 4; Length 507;  
 Best Local Similarity 98.0%; Pred. No. 6,7e-79;  
 Matches 440; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

QY 2660 ATATGACCATGTTTCAAAATTTATTAG-CCTATGAAAAATTTTCAGTGTGCACAT 2717  
 DB 59 AGACTCACCATGTTTCAAAATTTATTAGGCTTAATGAAAAATTTTCAGTGTGCACAT 118  
 QY 2718 CAATATTATTAGCTCAATTTACAGTTGATCTTCATTTCCATTAATAATACCCGAA 2777  
 DB 119 CAATATTATTAGCTCAATTTACAGTTGATCTTCATTTCCATTAATAATACCCGAA 178  
 QY 2778 TGTTCCTAAATTTACTTGTATGGGACAAATGTTAGTTCATTAATTTAGTCAAGGAC 2837  
 DB 179 TGTTCCTAAATTTACTTGTATGGGACAAATGTTAGTTCATTAATTTAGTCAAGGAC 238  
 QY 2838 TTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2897  
 DB 239 TTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 298  
 QY 2888 GATTTGTGAGTAT 2957  
 DB 299 GATTTGTGAGTAT 358  
 QY 2998 TGAATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3017  
 DB 359 TGAATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 418  
 QY 3018 ATTCATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3077  
 DB 419 ATTCATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 478  
 QY 3078 TTATTAGTTTCCCATTAATATGATTA 3106  
 DB 479 TTATTAGTTTCCCATTAATATGATTA 507

RESULT 3  
 US-08-998-416-66/c  
 Sequence 66, Application US/08998416  
 Patent No. 6239264  
 GENERAL INFORMATION:  
 APPLICANT: Philippen, Peter  
 APPLICANT: Pohlmann, Kainer  
 APPLICANT: Steiner, Sabine  
 APPLICANT: Mohr, Christine  
 APPLICANT: Wendland, Jürgen  
 APPLICANT: Knechtle, Philipp  
 APPLICANT: Reibischung, Corinne  
 TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPYII  
 TITLE OF INVENTION: AND USES THEREOF  
 NUMBER OF SEQUENCES: 1152  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: No. 6239264artis Corporation  
 STREET: 3054 Cornwallis Road  
 CITY: Research Triangle Park  
 STATE: No. 6239264th Carolina  
 COUNTRY: USA  
 ZIP: 27709  
 COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/998,416  
 FILING DATE: 24-DEC-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: CH 0016/97  
 FILING DATE: 31-DEC-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Meigs, J. Timothy  
 REGISTRATION NUMBER: 38,241  
 REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 919-541-8587  
 TELEFAX: 919-541-8689  
 INFORMATION FOR SEQ ID NO: 66:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 537 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 ORIGINAL SOURCE:  
 ORGANISM: PAG1020RP  
 US-08-998-416-66

Query Match 4.8%; Score 168.2; DB 3; Length 537;  
 Best Local Similarity 67.6%; Pred. No. 1.2e-25;  
 Matches 236; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 2610 CACGTCGAAGCTGATATTTGATATATTTAGTGATGATGATGATGATGATGATGATGAT 2669  
 DB 441 CTCAACCAAGAGCTGATATTTAGGCTCTATTTGGGATATTTAGGATGATGATGATGAT 382  
 QY 2670 TGTTCCTAAATTTACTTGTATGGGACAAATGTTAGTTCATTAATTTAGTCAATTTATTAG 2729  
 DB 381 TGTTCCTAAATTTACTTGTATGGGACAAATGTTAGTTCATTAATTTAGTCAATTTATTAG 322  
 QY 2730 CTCAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2789  
 DB 321 CACATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 262  
 QY 2790 TTACTTGTATGGGACATGTTAGTTCATTAATTTAGTCAATTTAGTCAATTTAGTCA 2849  
 DB 261 TCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 202  
 QY 2850 ATGTAGAGTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2909  
 DB 201 ACGTACCTGTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 142  
 QY 2910 TATTAATATTTATTAATTTATTTGATGATGATGATGATGATGATGATGATGATGAT 2958  
 DB 141 TTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 93

RESULT 4  
 US-09-248-796A-6302  
 Sequence 6302, Application US/09248796A  
 Patent No. 6747137  
 GENERAL INFORMATION:  
 APPLICANT: Keith Weinstock et al  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS  
 TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 107196.132  
 CURRENT APPLICATION NUMBER: US/09/248,796A  
 CURRENT FILING DATE: 1999-02-12  
 PRIOR APPLICATION NUMBER: US 60/074,725  
 PRIOR FILING DATE: 1998-02-13  
 PRIOR APPLICATION NUMBER: US 60/096,409  
 PRIOR FILING DATE: 1998-08-13



NUMBER OF SEQ ID NOS: 28208  
 SEQ ID NO 6302  
 LENGTH: 2445  
 TYPE: DNA  
 ORGANISM: Candida albicans  
 US-09-248-796A-6302

Query Match 3.4%; Score 118.2; DB 4; Length 2445;  
 Best Local Similarity 60.4%; Pred. No. 5.3e-15;  
 Matches 195; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 2624 GATATTCATTTATTTAGTGTATTTATTCAGATTCATATATACATGTTTCAAAATTTA 2683  
 DB 2101 GAAATTTGAAATGATTTAGTATATACAAAGATCATATGTTATCAATGTTGAGGCTTG 2160  
 QY 2684 TTAGCCTTAAAGAAAATTTTCAGTGTTCACATTCAAAATTTATTTAGTTCATTAAGTT 2743  
 DB 2161 GCCCATTTAGAAATGTTTACGAGATTCATTCGAATTTACTTGGCTCAATTAATATTT 2220  
 QY 2744 GAATCATTCATTCATTAATTAATTAATTCACCGAAATGTTTCTAAATTTACTTTGATTCG 2803  
 DB 2221 GATATGACCAAGATTAACACACATCAATGATGTTTGGAAATTCACATATTTGGGT 2280  
 QY 2804 ACAATGTAGTTCATTAATTTAGTCACGGGACTTTTGTATGATGTAAGAGTCCCT 2863  
 DB 2281 ACAAATGCTTGGCCCAATCATGTGTACTGGGCTCTGGGATGAACTGTTTGTCT 2340  
 QY 2864 GGTGAAGGTGTACCAATTTAGTGTGTTTTCGAAATTTGAGATTAATTAATTTA 2923  
 DB 2341 GGTCAAGATTAAGAGGCTTACCTGTTTGGGGAATTTGTTGATGATGCTTTGTTT 2400  
 QY 2924 ATTATTCATTCATTAATTTGCT 2946  
 DB 2401 AGCTTGTGCTTACATTTATGC 2423

## RESULT 5

US-09-248-796A-10622  
 Sequence 10622, Application US/09248796A  
 Patent No. 6747137

GENERAL INFORMATION:  
 APPLICANT: Keith Weinstock et al  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
 FILE REFERENCE: 107196.132  
 CURRENT APPLICATION NUMBER: US/09/248,796A  
 PRIOR FILING DATE: 1999-02-12  
 PRIOR APPLICATION NUMBER: US 60/074,725  
 PRIOR FILING DATE: 1998-02-13  
 PRIOR APPLICATION NUMBER: US 60/096,409  
 PRIOR FILING DATE: 1998-08-13  
 NUMBER OF SEQ ID NOS: 28208  
 SEQ ID NO 10622  
 LENGTH: 708  
 TYPE: DNA  
 ORGANISM: Candida albicans  
 US-09-248-796A-10622

Query Match 2.1%; Score 73.4; DB 4; Length 708;  
 Best Local Similarity 52.1%; Pred. No. 7.2e-06;  
 Matches 164; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 1130 GAACTGAAGAAGATGTTTCTTATGTTGTGATCATATTAAGATTAAGAAAT 1189  
 DB 385 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 444  
 QY 1190 GATTTTGAATGAATGAATGAATTTTAAAGAAAGAAAGAAAGCTTATTTCAAAAA 1249  
 DB 445 GAATTTGATGATGATGATGATGATTTTCAAGATGATGATGATGATGATGATGAT 504  
 QY 1250 CAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1309  
 DB 505 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 564

QY 1310 ACTACTAGTGTGATCTCTGTCATTCATATCATCATCATCACTAATTAATTAATTAAT 1369  
 DB 565 ACT 624  
 QY 1370 AATGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1429  
 DB 625 AACCTGATTAATTAATCATGCTGATTAATTTTAAATTCACACATCTTACGAAATTTGCT 684  
 QY 1430 TTAAGAAACATTA 1444  
 DB 685 GGGAATTTTAAACCTTA 699

## RESULT 6

US-08-487-826B-13  
 Sequence 13, Application US/08487826B  
 Patent No. 5993827

GENERAL INFORMATION:  
 APPLICANT: Sim, Kim L.  
 APPLICANT: Chitnis, Chetan  
 APPLICANT: Miller, Louis H.  
 APPLICANT: Peterson, David S.  
 APPLICANT: Su, Xin-zhuan  
 APPLICANT: Wellens, Thomas E.  
 TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
 TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
 NUMBER OF SEQUENCES: 45  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: Knobbe Martens Olson & Bear  
 STREET: 620 Newport Center Drive 16th Floor  
 CITY: Newport Beach  
 STATE: California  
 COUNTRY: US  
 ZIP: 92660  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/487,826B  
 FILING DATE: 10-SEP-1993  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Israelien, Ned  
 REGISTRATION NUMBER: 29,655  
 REFERENCE/DOCKET NUMBER: NIH121.001CP1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 235-8550  
 TELEFAX: (619) 235-0176  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 19124 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 HYPOTHEICAL: NO  
 ANTI-SENSE: NO  
 US-08-487-826B-13

Query Match 2.0%; Score 71.8; DB 2; Length 19124;  
 Best Local Similarity 46.8%; Pred. No. 5.1e-05;  
 Matches 290; Conservative 0; Mismatches 328; Indels 2; Gaps 2;

QY 2902 TGTGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2961  
 DB 826 TTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 885  
 QY 2962 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3021  
 DB 886 TTTACATGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 945

RESULT 7  
 US-09-806-708B-22  
 Sequence 22. Application US/09806708B  
 Patent No. 6784342  
 GENERAL INFORMATION:  
 APPLICANT: The University of British Columbia  
 TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants  
 FILE REFERENCE: 4810-58741  
 CURRENT FILING DATE: 2001-04-03  
 PRIOR APPLICATION NUMBER: US/09/806,708B  
 PRIOR FILING DATE: 1999-08-04  
 NUMBER OF SEQ ID NOS: 23  
 SOFTWARE: PatentIn version 3.0  
 SEQ ID NO 22  
 LENGTH: 1141  
 TYPE: DNA  
 ORGANISM: Artificial sequence  
 FEATURE:  
 NAME/KEY: Promoter  
 LOCATION: (1)..(1141)  
 OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. PAB1 promoters  
 US-09-806-708B-22

[illegible]

```

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12147
; LENGTH: 767677
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(767677)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12147

```

Query Match 1.9%; Score 68; DB 4; Length 767677;

Best Local Similarity 46.9%; Pred. No. 0.0012; Mismatches 240; Indels 0; Gaps 0;

```

Matches 212; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
QY 937 AACTAATGATAGTGAAGATATTACTTAATCTAGACCACTGCTAATCATATGAACTTGG 996
DB 343209 AGCCAAAGAAAACAAATATACCTTTCAACAGAACTTTTGAATGCAATCCTTTTAA 343268
QY 997 TATTGGTCTACACCCCTTGGTTGGAACTGGTACTACCGCACTGCCACTGCCACTGTC 1056
DB 343269 AAATGTTATGAGAGTGGTGAAGATGATGATGATGATGATGATGATGATGATGATG 343328
QY 1057 TGCTGCTGTGAAGAACCATCTCGTTCACTATTGATGATGATGATGATGATGATGATGATG 1116
DB 343329 TGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343388
QY 1117 AAGATCATCTCAAGAACTGAAGAGATGTTGTTTCTTATGTTGGTATCATATTTAG 1176
DB 343389 TGGTGGTATGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343448
QY 1177 AGTTAATGGAATGATTTTGTATGAAATTTGATGAATTTATTAGAGAGAAAGAAAGC 1236
DB 343449 TGATATATGAAGACGATGGTGGTGAAGATGATGATGATGATGATGATGATGATGATG 343508
QY 1237 TTATTACAAAACAAATGATGCTAAAATATTCGGGTATGATGATGATGATGATGATGATG 1296
DB 343509 TGATGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343568
QY 1297 TTCCAAAATATATCTACTAGTGTGATCTGTCATCCATATCATATCATATCAAGTAA 1356
DB 343569 TCATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343628
QY 1357 TAATTAATAAATAATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1388
DB 343629 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343660

```

RESULT 9

```

US-09-949-016-17361
; Sequence 17361, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

```

```

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12147
; LENGTH: 767677
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(767677)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17361

```

Query Match 1.9%; Score 68; DB 4; Length 767677;

Best Local Similarity 46.9%; Pred. No. 0.0012; Mismatches 240; Indels 0; Gaps 0;

```

Matches 212; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
QY 937 AACTAATGATAGTGAAGATATTACTTAATCTAGACCACTGCTAATCATATGAACTTGG 996
DB 343209 AGCCAAAGAAAACAAATATACCTTTCAACAGAACTTTTGAATGCAATCCTTTTAA 343268
QY 997 TATTGGTCTACACCCCTTGGTTGGAACTGGTACTACCGCACTGCCACTGCCACTGTC 1056
DB 343269 AAATGTTATGAGAGTGGTGAAGATGATGATGATGATGATGATGATGATGATGATG 343328
QY 1057 TGCTGCTGTGAAGAACCATCTCGTTCACTATTGATGATGATGATGATGATGATGATGATG 1116
DB 343329 TGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343388
QY 1117 AAGATCATCTCAAGAACTGAAGAGATGTTGTTTCTTATGTTGGTATCATATTTAG 1176
DB 343389 TGGTGGTATGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343448
QY 1177 AGTTAATGGAATGATTTTGTATGAAATTTGATGAATTTATTAGAGAGAAAGAAAGC 1236
DB 343449 TGATATATGAAGACGATGGTGGTGAAGATGATGATGATGATGATGATGATGATGATG 343508
QY 1237 TTATTACAAAACAAATGATGCTAAAATATTCGGGTATGATGATGATGATGATGATGATG 1296
DB 343509 TGATGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343568
QY 1297 TTCCAAAATATATCTACTAGTGTGATCTGTCATCCATATCATATCATATCAAGTAA 1356
DB 343569 TCATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343628
QY 1357 TAATTAATAAATAATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1388
DB 343629 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343660

```

RESULT 10

```

US-09-949-016-16284/C
; Sequence 16284, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 16284
; LENGTH: 29717

```

; TYPE: DNA  
 ; ORGANISM: Human  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)....(29717)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-949-016-16284

Query Match 1.9%; Score 67.6; DB 4; Length 29717;  
 Best Local Similarity 47.2%; Pred. No. 0.00044;  
 Matches 266; Conservative 0; Mismatches 295; Indels 2; Gaps 2;

QY 2960 AAAAAATGAAATATTCATTTAGAGACAAATAATGTAATGCAACATTTTAAATCAT 3019  
 DB 15962 AAGAACTTAATTTTAAAGTAAATAATATCATTTGTAGACAAATTCATTAATA 15903  
 QY 3020 TCATCAAGAGATCAATTGAAAGTTAGTTAAAAACATGCTGTATTAATCAATT 3079  
 DB 15902 TAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 15843  
 QY 3080 ATTAGTTCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3139  
 DB 15842 TGTAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 15783  
 QY 3140 TTTGTTTTTTTTTTTTTATGTCATGAGTTGTATATACATTAATCTTTTATAGAA 3199  
 DB 15782 AATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 15724  
 QY 3200 GTAACATATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3259  
 DB 15723 ATAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATGT 15664  
 QY 3260 GATATCTACTCTCTGATTTAGGAAAGTTATATTAATTAATTAATTAATTAAT 3319  
 DB 15663 AATAA-TAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 15605  
 QY 3320 TTTACGAGTTGTTAAATTTGAGAGTCATTAATTAATTAATTAATTAATTAATG 3379  
 DB 15604 AATATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 15545  
 QY 3380 AAGAAATTAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3439  
 DB 15544 TATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 15485  
 QY 3440 GATATATCGCATTTGTTGTAATTTTGAATTAATTAATTAATTAATTAATTA 3499  
 DB 15484 AATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 15425  
 QY 3500 TTATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3522  
 DB 15424 TAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 15402

RESULT 11  
 US-09-949-016-12423/c  
 ; Sequence 12423, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: C1001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 12423

; LENGTH: 60376  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)....(60376)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-949-016-12423

Query Match 1.9%; Score 67.6; DB 4; Length 60376;  
 Best Local Similarity 47.2%; Pred. No. 0.00057;  
 Matches 266; Conservative 0; Mismatches 295; Indels 2; Gaps 2;

QY 2960 AAAAAATGAAATATTCATTTAGAGACAAATAATGTAATGCAACATTTTAAATCAT 3019  
 DB 48001 AAGAACTTAATTTTAAAGTAAATAATATCATTTGTAGACAAATTCATTAATA 47942  
 QY 3020 TCATCAAGAGATCAATTGAAAGTTAGTTAAAAACATGCTGTATTAATCAATT 3079  
 DB 47941 TAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 47882  
 QY 3080 ATTAGTTCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3139  
 DB 47881 TGTAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 47822  
 QY 3140 TTTGTTTTTTTTTTTTTATGTCATGAGTTGTATATACATTAATCTTTTATAGAA 3199  
 DB 47821 AATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 47763  
 QY 3200 GTAACATATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3259  
 DB 47762 ATAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATGT 47703  
 QY 3260 GATATCTACTCTCTGATTTAGGAAAGTTATATTAATTAATTAATTAATTAAT 3319  
 DB 47702 AATAA-TAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 47644  
 QY 3320 TTTACGAGTTGTTAAATTTGAGAGTCATTAATTAATTAATTAATTAATTAATG 3379  
 DB 47643 AATATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 47584  
 QY 3380 AAGAAATTAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3439  
 DB 47583 TATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 47524  
 QY 3440 GATATATCGCATTTGTTGTAATTTTGAATTAATTAATTAATTAATTAATTA 3499  
 DB 47523 AATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 47464  
 QY 3500 TTATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3522  
 DB 47463 TAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 47441

RESULT 12  
 US-09-601-198-56/c  
 ; Sequence 56, Application US/09601198  
 ; Patent No. 653183  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cassell, Gail H.  
 ; APPLICANT: Chen, Ellison Y.  
 ; APPLICANT: Glass, Jennifer S.  
 ; APPLICANT: Heiner, Cheryl R.  
 ; APPLICANT: Lefkowitz, Elliot  
 ; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREA PLASMA  
 ; TITLE OF INVENTION: UREA PLASMA  
 ; FILE REFERENCE: UAB-13452/22  
 ; CURRENT APPLICATION NUMBER: US/09/601,198  
 ; CURRENT FILING DATE: 2000-12-08  
 ; PRIOR APPLICATION NUMBER: 60/073,189  
 ; PRIOR FILING DATE: 1998-01-30  
 ; NUMBER OF SEQ ID NOS: 181



[illegible]

```

RESULT 14
US-09-843-376-10/c
; Sequence 10, Application US/09843376
; Patent No. 6566132
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF INTERFERON GAMMA RECEPTOR 1 EXPRESSION
; FILE REFERENCE: RRS-0234
; CURRENT APPLICATION NUMBER: US/09/843,376
; NUMBER OF FILING DATE: 2001-04-26
; NUMBER OF SEQ ID NOS: 88
; SEQ ID NO 10
; LENGTH: 26000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-843-376-10

Query Match      1.9%; Score 66; DB 4; Length 26000;
Best Local Similarity 48.0%; Pred. No. 0.00091;
Matches 225; Conservative 0; Mismatches 236; Indels 8; Gaps 1;

QY 3055 AAAACATGCTGGTATTAATCAATTATTAGTTCCCAATAAATATGAAATAGCAATATC 3114
Db 24147 AATCTAATATATCTAAATAAACCTCCCTTATATATATATATATATATATATTA 24088
QY 3115 AAAGAAATGCACAGAGTTGATGGTTGTTTTTTTTTTTTTTTTTTTTTATGTCATGAGT 3174
Db 24087 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 24028
QY 3175 GTATATACATATACTTTTTTATAGAAAGTAACAATATGTAATATGATATAGTCAATC 3234
Db 24027 ATTATATAAATATATATATATATATATATATATATATATATATATATATATAT 23968
QY 3235 ATCATATTTATATATGATATATATATATATATATATATATATATATATATATATAT 3294
Db 23967 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 23908
QY 3295 ATATTATTTACTATATAACATTATTTTTTATAGAGTGTGTTAAATATGAGTCAATTTA 3354
Db 23907 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 23856
QY 3355 TAGATGTAAAGAAGTTTATTAAGAGAAATTAATATATATATATATATATATATATAT 3414

```

[illegible]

```

US-09-949-016-15851/c
; Sequence 15851, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01037
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 15851
; LENGTH: 205044
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(205044)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15851

```

Query Match	Similarity	1.9%	Score 65.8	DB 4	Length 205044
Best Local	Similarity	44.3%	Pred. No. 0.0021		
Matches	312	Conservative	0	Mismatches 332	Indels 1
				Gaps	1
Qy	2631	CATTATATTTAGTGATATTCACAGATCATATATACACAGTGTCCAAATTTATTAGCCT	2650		
Db	201143	CATTAAATATATTCATCAATATTTATATTTATTTATTTAAATATATATATTTTAT	2010		
Qy	2691	ATGMAAAATTTTCAGTCGTTCCATTCGAATTTATAGTCATTTCAAGTTCAATCAT	2750		
Db	201083	ATATTTATATTTAAATATATATATATTTTATATATATATTTTAAATATATATATTTT	2010		
Qy	2751	TCAATTCGAATATATATATCAACCGAAATGTTTTCTAAATATACCTTGATGGGCAATG	2810		
Db	201023	ATATATTTATATTTAAATATATATATATTTTATATATATTTTAAATATATATATTT	2009		
Qy	2811	TAGTTCATTTAAATTTAGTCACGGGACTTTTGGTATGAATGTAAAGTCCCTGGTAG	2870		
Db	200963	TTATATATTTTATATTTTAAATATATATATATTTTATATATTTTAAATATATATAT	2009		
Qy	2871	GTGGTACCAATTTAGTGTGGTTTTCCGAATGTGTGAGATTTATATTTTATATTTATG	2930		
Db	200903	TTTATATATTTTATATTTTAAATATATATATTTTATATTTTATATTTTAAATATATAT	2008		
Qy	2931	GATCATTTATTTTGTCTCATGTGGTGTGAAAAATTTGAATATTCATTCAGAGACAA	2990		
Db	200843	ATTTTATATATTTTATATTTTAAATATATATATTTTATATTTTATATTTTAAATATAT	2007		
Qy	2991	ATAATGTAAATCGACCAATTTTATCATTCATCAAGACATCAATAGAGTTAGGTT	3050		
Db	200783	ATAATATATTTTATATATTTTATATTTTAAATATATATATATATTTTATATATTTATA	2007		
Qy	3051	TAAAAAAACATGTGGTATATATCAATTTATAGTTCCCAATATATATGATATAGCAT	3110		

